



# Etude de la régulation de l'expression des gènes par un ARN antisens

Stanislas Denoeux

## ► To cite this version:

Stanislas Denoeux. Etude de la régulation de l'expression des gènes par un ARN antisens. Biologie cellulaire. Université Paris Saclay (COMUE), 2015. Français. NNT : 2015SACLS167 . tel-01276303

**HAL Id: tel-01276303**

**<https://theses.hal.science/tel-01276303>**

Submitted on 19 Feb 2016

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**NNT 2015SACLS167**

**université**  
**PARIS-SACLAY**

ÉCOLE DOCTORALE  
**Cancérologie, biologie,  
médecine, santé (CBMS)**

**THESE DE DOCTORAT**  
**DE**  
**L'UNIVERSITE PARIS-SACLAY**  
**PREPAREE A**  
**L'UNIVERSITE PARIS-SUD**

ÉCOLE DOCTORALE N°582  
Cancérologie, biologie, médecine, santé

Aspects moléculaires et cellulaires de la biologie

Par

**M. Stanislas DENOEU**

**Etude de la régulation de l'expression des gènes par les ARN antisens**

**Thèse présentée et soutenue à Cachan, le 14 décembre 2015 :**

**Composition du Jury :**

Pr. Yannick ANDEOL	Rapporteur
Dr. Christophe ANTONIEWSKI	Rapporteur
Pr. Karim BENIHOUD	Président du jury / Examineur
Dr. Claude BOUCHEIX	Examineur
Dr. Martin DUTERTRE	Examineur
Pr. François DAUTRY	Directeur de thèse

## Remerciements

Je voudrais en premier lieu remercier François de m'avoir accueilli dans son équipe et de m'avoir proposé ce sujet de thèse. Je le remercie pour ses nombreux conseils, son encadrement attentif et son soutien. Je suis fier et reconnaissant d'avoir pu travailler avec quelqu'un de son expérience pendant ces trois années et je tiens tout particulièrement à le remercier de son aide indispensable durant la rédaction de ce manuscrit.

J'aimerais également remercier sincèrement l'ensemble des membres du jury d'avoir accepté mon invitation à examiner mon travail. Mes remerciements s'adressent en particulier au Pr. Yannick ANDEOL et Dr. Christophe ANTONIEWSKI, qui ont consacré leur temps et leur énergie pour évaluer ce manuscrit et m'ont apporté des commentaires pertinents.

J'adresse de la même façon mes remerciements à mes collègues de l'équipe, Ilham, Maryline, Anh Thu, Lue et Thierry qui m'ont non seulement beaucoup aidé pour mes expériences, mais m'ont aussi beaucoup soutenu psychologiquement tout au long de mes travaux de recherche. Je remercie également le LBPA et l'ensemble de ses membres de m'avoir accueilli chaleureusement pendant ma thèse.

Mes remerciements vont aussi à ma famille et mes proches qui, malgré leurs difficultés à comprendre réellement ce que je faisais, m'ont toujours soutenu. Je les remercie pour leurs questions comme « Tu en es où ? » ou « Tu finis quand ta thèse ? », qui bien qu'angoissante m'ont permis de ne jamais dévier de mon objectif final. Un grand merci surtout à mes parents qui ont toujours été là pour moi durant ces très longues études.

Enfin, mon dernier remerciement et non des moindres va à celle avec qui je partage ma vie et qui a dû me supporter dans mes bons et surtout mes mauvais jours. Je te remercie Hai-mi d'avoir toujours su trouver les mots et les gestes pour me reconforter et me remotiver dans ces moments difficiles.



**Titre :** Etude de la régulation de l'expression des gènes par les ARN antisens

**Mots clés :** ARN, ARN antisens, Régulation génétique

**Résumé :** Au cours de la dernière décennie, les avancées du séquençage à haut débit ont permis de caractériser un grand nombre d'ARN non codants et d'établir l'existence de transcrits "antisens" pour de nombreux gènes chez les mammifères. Cependant leur rôle dans le contrôle de l'expression des gènes "sens" auxquels ils sont associés est encore très mal connu.

Mes travaux ont porté sur la caractérisation de certains aspects du mécanisme d'action des longs ARN non codants. Ils reposent sur le développement d'une approche de constructions indicatrices fluorescentes dont l'expression est suivie par cytométrie en flux en présence ou non d'ARN "antisens". Cette approche a le potentiel de mettre en évidence une régulation même si elle n'est présente que dans une sous-population cellulaire.

Une première série d'expériences a été réalisée en expression transitoire pour bénéficier d'un contexte chromatinien simplifié. Mais dans ce cas, les silencing observés sont aussi actifs sur une construction contrôle, indiquant la mise en place d'une réponse non spécifique de séquence qui évoque la réponse de type interféron. Cependant, l'expression globale des gènes cellulaires n'est pas significativement affectée, indiquant une certaine spécificité de la réponse. Parmi les voies testées, ni la kinase PKR, ni la RNaseL ou la voie de l'interférence par l'ARN ne peuvent rendre compte du silencing observé. Une des caractéristiques de cette régulation est qu'elle n'affecte pas les gènes intégrés au génome mais uniquement les gènes exprimés à partir d'une construction épisomale ce qui évoque des caractéristiques souhaitables pour un mécanisme antiviral. Cependant l'ampleur de cette réponse non spécifique empêche toute étude plus approfondie d'un mécanisme spécifique s'il existe.

Mes travaux se sont alors portés sur l'étude de ces mêmes constructions en clone stable dans deux contextes différents pour l'expression de l'ARN antisens : en *cis* ou en *trans*.

Dans le cas de l'expression en *trans*, un ARN antisens sans séquence régulatrice particulière ne permet pas la mise en place d'un silencing. Cette observation est en accord avec le faible nombre de longs ARN antisens connus pour agir en *trans* dans la nature.

Par contre l'expression en *cis* d'un ARN antisens peut conduire à un silencing spécifique. Cette organisation dans laquelle les gènes « sens » et « antisens » sont situés sur le même fragment d'ADN correspond à celle majoritairement observée pour les longs ARNs antisens dans la nature (*cis*NAT, *cis* Natural Antisense Transcripts). Cependant, mes travaux montrent que le silencing observé n'est pas stable dans le temps et disparaît dès lors que la transcription antisens cesse, indiquant l'absence d'une mémoire épigénétique. Un tel mode de régulation est compatible avec une interférence transcriptionnelle dans laquelle la transcription et non le produit ARN est la cause du silencing. Par ailleurs, un certain nombre de cas de co-régulation du transcrit sens et antisens ont été observés, ce qui traduit la possibilité d'activer en *cis* la transcription du gène cible par le promoteur de son ARN antisens. Ce phénomène est probablement facilité par la petite taille des constructions, mais cette dualité de réponse est en accord avec l'absence de corrélation (positive ou négative) entre l'expression des gènes et de leur transcrits antisens. L'ensemble de mes travaux montrent la faible capacité d'un ARN antisens à induire un silencing. L'approche développée doit donc permettre de rechercher des co-activateurs du silencing, par exemple en introduisant des sites de recrutement de complexes modificateurs de la chromatine.



**Title :** Regulation of gene expression by antisense RNA

**Keywords :** RNA, RNA antisense, Gene silencing

**Abstract :** During the last decade, next generation sequencing has allowed the characterization of a large number of non-coding RNA and the establishment of the fact that a majority of mammalian genes were also transcribed in the opposite orientation. However the functional significance of this antisense transcription is currently unclear.

My work focused on the characterization of the regulatory potential of long non-coding RNA. It relied on the use of fluorescent reporter constructs. Their expression in the presence or absence of antisense RNA is analyzed by flow cytometry. This approach has the potential to uncover a regulation mechanism even if it takes place only in a subpopulation of cells.

A first series of experiments has been performed by transient expression assays in order to benefit from a simplified chromatin context. In this case the silencing associated with antisense transcripts is also active on control constructs, indicating that at least part of the response is not sequence specific suggesting the involvement of an interferon-type response.

However, cellular gene expression is not significantly affected indicating some level of specificity. Among the investigated pathways, neither the PKR kinase, nor RNaseL or RNA interference pathway can account for the observed silencing. One of this regulation attributes is that it does not affect genes integrated in the genome but only genes expressed from episomes, a selectivity which would seem appropriate for an antiviral mechanism. Nevertheless the extent of this non-specific response impedes any further study of a specific mechanism, if it operates.

My work focused on the study of these reporter constructs after their integration in the genome, with antisense RNA being expressed in *cis* or in *trans*.

In the case of *trans* expression, an antisense RNA devoid of any specific regulatory sequence does not allow the setting of a silencing. This observation is consistent with the low number of long antisense RNA known to naturally act in *trans*.

On the other side, the *cis* expression of an antisense RNA can lead to a specific silencing. This organization in which “sense” and “antisense” genes are located in the same DNA fragment matches with the ones mostly observed for long antisense RNA in nature (*cis*NAT, *cis* Natural Antisense Transcripts).

However, my work shows that the observed silencing is not stable over time and the effects terminate once antisense transcription stops, indicating the absence of an epigenetic memory. This regulation mode is consistent with a transcriptional interference in which transcription – and not its RNA product - is causing the silencing. Besides, I observed a certain number of sense and antisense transcript co-regulation cases highlighting the possibility to activate the transcription of the target gene by the promoter of its antisense RNA. This phenomenon is probably facilitated by the small size of the constructs, but this duality of response is in agreement with the lack of correlation (either positive or negative) between the expression of genes and their antisense transcripts.

This study shows the limited capacity of an antisense RNA to induce a silencing. The developed approach should allow the search for silencing co-activators, for instance by introducing chromatin remodeling complexes recruitment sites.



<b>INTRODUCTION.....</b>	<b>1</b>
I - LES ARN NON CODANTS.....	1
1 – <i>Les petits ARN régulateurs</i> .....	2
1.1 – Découverte et première description .....	2
1.2 – L’interférence par les siARN .....	5
1.3 – L’interférence par les miARN .....	8
1.4 – L’interférence par les piARN.....	15
2 – <i>Les longs ARN non codants (lncRNA)</i> .....	20
2.1 – Première description et caractéristiques .....	20
2.2 – Une multiplicité de modes d’action .....	23
3 – <i>Les longs ARN antisens</i> .....	40
3.1 – Généralités et spécificités .....	40
3.2 – Mode d’action : différence et similarité avec les longs ARN non codants .....	41
II – LES REGULATIONS CYTOPLASMIQUES EN REPONSE A LA PRESENCE D’ARN DB .....	49
1 – <i>PKR – Protein Kinase RNA-activated</i> .....	49
2 – <i>OAS – OligoAdenylate Synthase</i> .....	50
3 – <i>RIG-I – Retinoic acid Inducible Gene I</i> .....	50
4 – <i>MDA5 – Melanoma Differentiation Associated gene 5</i> .....	51
5 – <i>TLR3 – Toll-like receptor 3</i> .....	52
6 – <i>Les protéines ADAR – L’édition ARN</i> .....	53
III – PROBLEMATIQUE .....	55
<b>MATERIELS ET METHODES.....</b>	<b>58</b>
I – CELLULES .....	58
1 – <i>La lignée cellulaire</i> .....	58
2 – <i>Culture cellulaire</i> .....	59
3 – <i>Transfection</i> .....	59
3 – <i>Construction des clones stables HEK 293 Tet-on</i> .....	61
II – CYTOMETRIE EN FLUX.....	61
1 – <i>Préparation des échantillons</i> .....	61
2 – <i>Protocole d’analyse</i> .....	62
2.1 – Analyse en deux couleurs : GFP et DsRed .....	62
2.2 – Analyse en trois couleurs : GFP, DsRed et Nir .....	64
III – BIOLOGIE MOLECULAIRE .....	64
1 – <i>Les plasmides</i> .....	64
1.1 – Les constructions pBiFluo et dérivés .....	65
1.2 – Les constructions indicatrices CMV .....	67



1.3 – La construction pStan .....	68
1.4 – Le Binding Element PRC2.....	69
<b>2 - Les siARNs .....</b>	<b>69</b>
<b>3 - Extraction des ARNs et RT-qPCR.....</b>	<b>70</b>
3.1– Extraction des ARNs .....	70
3.2 – Reverse transcriptase .....	71
3.3 – Analyse par qPCR.....	71
<b>RESULTATS.....</b>	<b>73</b>
<b>1 – EFFET EN TRANSITOIRE .....</b>	<b>73</b>
1.1– Première observation de l'effet d'un antisens.....	73
1.3 – Un effet transcriptionnel de l'antisens ? .....	75
1.4 – Etude de la réponse non spécifique PKR-dépendante .....	76
1.4.1 – Inhibition de PKR par l'inhibiteur chimique C16 .....	76
1.4.2 – Inhibition de PKR par l'utilisation d'un siARN .....	79
1.5 – Etude de l'effet d'un ARN antisens exprimé en cis .....	82
1.6 – Implication de l'ARN interférence ? .....	86
<b>2 – EFFET EN CLONES STABLES .....</b>	<b>88</b>
2.1 –Silencing en trans d'une construction intégrée .....	90
2.1.1 – L'effet d'un antisens exprimé en transitoire .....	90
2.1.2 – L'effet d'un antisens exprimé en intégré .....	94
2.1.3 – Effet de l'ajout d'une séquence de recrutement .....	99
2.1.4 – Effet de l'absence de séquence de polyadénylation.....	100
2.1.5 – Effet de l'inhibition de PAF1.....	104
2.2 –Silencing en cis d'une construction intégrée .....	105
2.2.1 – Analyse du silencing dans la construction pStan intégrée.....	105
2.2.2 – Etude de l'insertion d'une séquence PolyA dans pStan.....	110
2.2.3 – Ajout de l'élément PRC2 à la construction pStan intégrée.....	113
<b>DISCUSSION .....</b>	<b>116</b>
1 – LA REPONSE EN TRANSITOIRE ET LES EFFETS NON SPECIFIQUES .....	116
2 – L'ETUDE DE L'EFFET DE L'ARN ANTISENS EN CLONES STABLES.....	122
3 - BILAN ET PERSPECTIVES.....	130
<b>BIBLIOGRAPHIE .....</b>	<b>133</b>



## INTRODUCTION

### I - Les ARN non codants

Depuis les années 1950, la biologie moléculaire s'est construite autour d'une chaîne d'événements censée décrire le lien entre l'information génétique et le fonctionnement des systèmes biologiques : l'ADN est transcrit en ARN qui est traduit en protéines, ces dernières portant l'ensemble des activités biologiques opérationnelles de la cellule. Ce schéma qui se concentre sur les ARN messagers (ARNm) a été construit alors même que les premiers ARN non codants étaient décrits. C'est en particulier le cas des ARN ribosomiques (ARNr) qui représentent près de 80% du transcriptome. La famille des ARNr est composée de 4 ARN chez les eucaryotes qui sont les ARN 28 S, 5,8 S et 5 S qui forment l'ossature sur laquelle s'assemble la grande sous-unité ribosomique 60 S et l'ARN 18 S sur lequel s'assemble la petite sous-unité 40 S. Si le rôle attribué aux ARNr a longtemps été vu comme purement structural, il est apparu plus récemment avec les premières données cristallographiques sur le ribosome obtenues en 2000 que le cœur catalytique du ribosome qui assure la formation de la liaison peptidique est en fait porté par l'ARN (Moore *et al*, 2011). Ainsi, non seulement, la classe la plus abondante d'ARN ne code pas une protéine mais elle fournit aussi l'activité catalytique au cœur de la traduction en accord avec l'idée d'un monde primordial constitué d'ARN. De plus, avant même les ARNr, Hoagland et Zamecnik avait décrit en 1958 les ARN de transfert (ARNt) qui permettent la lecture du code génétique. Leur structure secondaire est formée de 3 tiges boucles qui permettent la reconnaissance du codon de l'ARNm et le transfert de l'acide aminé correspondant à la protéine en cours de synthèse au sein du ribosome. Les ARNt sont associés aux acides aminés par une liaison ester à leur extrémité 3', catalysée par l'aminoacyl-ARNt synthétase. Les ARNr et ARNt ne sont d'ailleurs que les premiers d'une longue liste d'ARN non codant découverts par la suite comme les petits ARN nucléaires (ou snRNA pour « small nuclear RNA ») impliqués dans la machinerie d'épissage ou les petits ARN nucléolaires (ou snoARN) qui participent à la maturation des ARNr. La découverte de ces ARN indispensables à la maturation des transcrits et à la synthèse protéique n'a cependant pas suffi à changer le point de vue sur le rôle des ARN non codants, leurs fonctions étant perçues comme essentiellement « mécaniques » par





opposition avec les fonctions « nobles » de régulation assurées par les protéines. Pourtant dès 1989, il avait été observé que les ARN, davantage que les protéines, sont indispensables à l'organisation du noyau (Nickerson *et al*, 1989), mais là encore ceci pouvait être interprété comme l'indication d'un rôle structurel plus que régulateur.

Au-delà des ARNr et des différents types de petits ARN, qui pour l'essentiel sont transcrits par des ARN polymérases spécialisées (Pol I et Pol III), des ARN non codants d'une taille et d'une organisation similaire à celle des ARN messagers, transcrits par la Pol II, ont progressivement été décrits. Le fait qu'ils étaient souvent associés à des phénomènes de suppression de l'expression, comme Xist pour l'inactivation de l'X ou H19 pour l'empreinte parentale, suggérait un rôle régulateur au moins par défaut, tant que l'on ne parvenait pas à détecter une protéine. Ce n'est donc qu'avec le NGS (« Next Generation Sequencing ») et le développement du séquençage des ARN que la question des ARN non codants a pris toute son ampleur. En effet, si les premières indications d'une transcription de l'ensemble du génome ont été obtenues par des études sur puces, celles-ci pouvaient être faussées par de nombreux artefacts d'hybridation. Par contre, le séquençage à haut débit a clairement établi que la grande majorité du transcriptome des eucaryotes supérieurs ne code aucune protéine et est donc constitué d'ARN non codants. Pour autant, ces observations ne permettent pas de clore le débat puisque le rôle biologique de la plupart de ces transcrits reste encore à établir. Il a d'ailleurs été proposé qu'une majorité de ces transcrits n'ait pas vraiment de fonction biologique et que leur existence soit due à l'impossibilité de réaliser un contrôle strict de la transcription à l'échelle du génome complet (Struhl *et al*, 2007). L'importance de cette question de la fonction des ARN non codants fait que la présentation qui va suivre est organisée autour des fonctions plutôt que des transcrits eux-mêmes.

## 1 – Les petits ARN régulateurs

### 1.1 – Découverte et première description

Si les ARN non codants ont des fonctions diverses, un groupe particulier semble se dégager, celui des ARN régulateurs. Ce sont des ARN de taille variable qui sont impliqués dans la régulation de l'expression des gènes. Leur taille va déterminer les mécanismes impliqués. Les premiers à avoir été décrits sont les petits ARN participant à « l'ARN interférence ». L'introduction du terme « ARN interférence » vient des travaux de Fire et Melo (Fire *et al*,



# Introduction

---

1998) qui ont mis en évidence chez *C. elegans* un mécanisme de régulation de l'expression des gènes par des ARN double brin. Les auteurs ont montré que l'introduction par micro-injection dans les cellules de *C. elegans* d'ARN double brin induit une inhibition de l'expression de tout gène qui contient une région identique en séquence avec celle des ARN double brin. Ils ont aussi établi que ce mécanisme est post-transcriptionnel et affecte l'accumulation des ARN messagers correspondants. Ainsi la micro-injection dans quelques cellules de *C. elegans* est suffisante pour induire un « silencing » aussi appelé PTGS (« Post-Transcriptional Gene Silencing ») qui se réfère à un groupe de mécanismes par lequel l'expression d'un ou plusieurs gènes est régulée négativement ou entièrement supprimée par l'introduction d'un ARN antisens.

Ces observations où l'ARN semble être l'acteur principal d'une régulation génique constituaient clairement une rupture des dogmes de la biologie moléculaire. Rapidement après la description initiale chez *C. elegans*, ce mécanisme a été mis en évidence chez d'autres espèces comme la drosophile (Kennerdell et Carthew, 1998) ou la souris (Svoboda *et al*, 2000). De fait, il fut démontré que l'interférence par l'ARN est un mécanisme conservé chez la grande majorité des eucaryotes et qu'il existait un mécanisme au moins partiellement apparenté chez les bactéries ou chez les archaebactéries suggérant une origine très ancienne de ce type de mécanisme. Sa découverte chez *C. elegans* avait d'ailleurs été précédée par des descriptions de phénomènes non mendéliens qui suggéraient l'existence de mécanismes pas encore identifiés. C'est le cas en particulier chez les plantes, où le phénomène de co-suppression avait été décrit une dizaine d'années auparavant par Jorgensen et ses collaborateurs (Napoli *et al*, 1990). L'objectif de leurs travaux était de renforcer la couleur pourpre des pétunias en introduisant des copies supplémentaires du gène codant l'enzyme clé de la synthèse de ce pigment (la chalcone synthase). L'introduction d'une copie supplémentaire du gène conduit fréquemment à l'extinction du gène cellulaire donnant un phénotype partiellement ou totalement blanc aux pétunias. L'ensemble des données montre que ce mécanisme semble avoir été conservé au cours de l'évolution pour lutter contre les virus et plus spécialement les virus à ARN. Chez les plantes, par exemple, il existe un grand nombre de virus qui possèdent un génome ARN et l'inactivation des gènes impliqués dans l'ARN interférence va conduire à des infections beaucoup plus sévères (Vaucheret *et al*, 2001). De plus, il a été montré qu'un grand nombre de virus de plantes



# Introduction

---

étaient capables « d'interférer » avec l'ARN interférence, signe que ce dernier exerce une pression de sélection sur les virus (Ding et Voinnet, 2007). Il a d'ailleurs été proposé que l'absence de la machinerie de l'interférence ARN, observée chez *S. cerevisiae* et quelques autres levures, reflète sa perte du fait d'une pression de sélection favorable au maintien du « Killer virus » (Drinnenberg *et al*, 2011).

Dans un premier temps, l'utilisation de l'ARN interférence chez les mammifères a été considérée comme impossible du fait de l'existence de la réponse interféron. En effet, chez les mammifères l'introduction d'ARN double brin induit des réponses cellulaires dont la production d'interférons qui bloquent la synthèse protéique de façon indépendante de la séquence de l'ARN antisens. Mais, rapidement après la description initiale de l'interférence, il a été mis en évidence la présence de petits ARN d'une vingtaine de nucléotides (siARN pour « short interfering RNA ») chez les plantes où était observée la co-suppression (Hamilton et Baulcombe, 1999) ce qui a rendu possible l'utilisation de cette technique chez les mammifères. En effet, la description d'une nucléase spécifique de séquence induite par les ARN double brin chez la drosophile (Hammond *et al*, 2000), suggérait que ces petits ARN interférents pouvaient jouer le rôle de guide dans la reconnaissance des ARN cibles. Ceci a donné l'idée à l'équipe de Thomas Tuschl d'utiliser directement les petits ARN produits du clivage des longs ARN double brin pour induire une interférence ARN chez les mammifères (Elbashir *et al*, 2001). Depuis, une meilleure compréhension du mécanisme de l'ARN interférence en a fait un puissant outil de biologie moléculaire pour le « Knock Down » ou mise sous silence (assourdissement) des gènes afin d'étudier leur fonction.

Néanmoins, il a été observé très vite que les siARN n'étaient pas les seuls petits ARN capables de faire de « l'interférence ». En 2001, plusieurs équipes dont celle de Thomas Tuschl identifient un autre type de petits ARN, issus d'ARN précurseurs simple brin mais qui contiennent une région fortement appariée dans une structure secondaire en tige-boucle et qui donnent naissance aux « micro-ARN » ou miARN (Lagos-Quintana *et al*, 2001). En fait, ce type d'ARN régulateur avait été observé dès 1993 chez *C. elegans* avec la caractérisation du produit du gène lin-4 (Lee *et al*, 1993), mais jusqu'alors il avait été supposé que c'était une spécificité de *C. elegans*. Puis en 2006, ce sont les « piwi interacting RNA » ou piARN qui sont découverts chez la drosophile, leur fonction est de protéger la lignée germinale contre les éléments transposables (Aravin *et al*, 2006 ; Girard *et al*, 2006) établissant l'idée qu'il existe



plusieurs classes d'interférence (Aravin *et al*, 2001). Ces 3 types de petits ARN diffèrent dans leur biogénèse et leur mode d'action que je vais détailler.

## 1.2 – L'interférence par les siARN

Le silencing par les siARN a été le premier à être décrit. Celui-ci est impliqué dans la réponse antivirale de la cellule induite par la présence d'ARN double brin dans le cytoplasme. Il est aujourd'hui utilisé comme un outil moléculaire pour le silencing de l'expression des gènes dans les études fonctionnelles.

Comme pour les autres petits ARN régulateurs, le mécanisme d'action de ces petits ARN va reposer sur sa capacité à interagir avec un complexe de protéines régulatrices et à le guider vers un ARN « cible ». La première découverte faite dans ce sens a été l'observation d'une activité nucléase spécifique de séquence induite par l'introduction d'ARN double brin chez la drosophile (Hammond *et al*, 2000) et baptisée RISC pour « RNA Induced Silencing Complex ». Très vite un petit ARN d'une vingtaine de nucléotides a été décrit en interaction avec cette nucléase lui permettant de la guider vers l'ARN messenger cible. L'action catalytique de cette nucléase, de la famille des protéines Argonaute, est au cœur du mécanisme de l'interférence par les siARN. La mise en évidence de la protéine Argonaute a été faite pour la première fois chez *A. thaliana* dans l'étude d'une mutation affectant la forme des feuilles (Bohmert *et al*, 1998). Il fut montré que la protéine Argonaute 1 était nécessaire au mécanisme de co-suppression chez *A. thaliana* et que le gène codant cette protéine était apparenté au gène Rde1 (« RNAi deficient ») de *C. elegans*. Le groupe d'Hannon a étendu ces études à la drosophile en impliquant la protéine Argonaute 2 (Ago2) dans le silencing post-transcriptionnel. Depuis, il a été montré qu'il existait des protéines de la famille Argonaute chez tous les eucaryotes, à l'exception déjà mentionnée de quelques levures dont *S. cerevisiae*. L'étude phylogénétique de la famille des protéines Argonaute (Parker et Barford, 2006) montre une grande diversité de protéines Argonaute chez *C. elegans* et les plantes, reflétant l'importance biologique de ce type de régulation chez ces organismes. Chez les mammifères, il existe 7 à 8 protéines Argonaute qui peuvent être regroupées en deux sous-familles : les protéines Argonaute (Ago1, 2, 3 et 4) et les protéines Piwi (Piwi-Like 1, 2 et 3 chez la souris). La sous-famille des protéines Piwi est absente chez les plantes, les



# Introduction

---

champignons et *S. pombe* et a donc une distribution plus restreinte que celle des protéines de la sous-famille Argonaute. Elles sont impliquées dans l'interférence par les piARN (décrite plus loin) qui est essentielle pour la régulation des éléments transposables dans les cellules de la lignée germinale.

Le complexe RISC est donc organisé autour d'une protéine Argonaute et d'un petit ARN guide. Chez les mammifères, seule Ago2 a une activité endonucléolytique qui permet le clivage de l'ARN cible. Cette activité est portée par le domaine Piwi de la protéine comme le montre la structure de la protéine Argonaute de *P. furiosus* (Song *et al*, 2004). Cette étude a mis en évidence une homologie du domaine Piwi avec le domaine catalytique des RNase H des rétrovirus des vertébrés capable d'hydrolyser le brin ARN dans un double brin ADN:ARN. Dans le cas de la protéine Ago2, l'hydrolyse de l'ARN ciblé ne demande pas de caractéristique particulière mise à part la complémentarité parfaite avec le petit ARN « guide » (ou presque parfaite, un ou deux mésappariements en dehors du site de coupure pouvant être tolérés). Cela permet d'ailleurs d'utiliser l'interférence ARN sur des transcrits non codants ou sans queue polyA tandis que la traduction n'est pas déterminante dans l'efficacité du silencing (Sen *et al*, 2005).

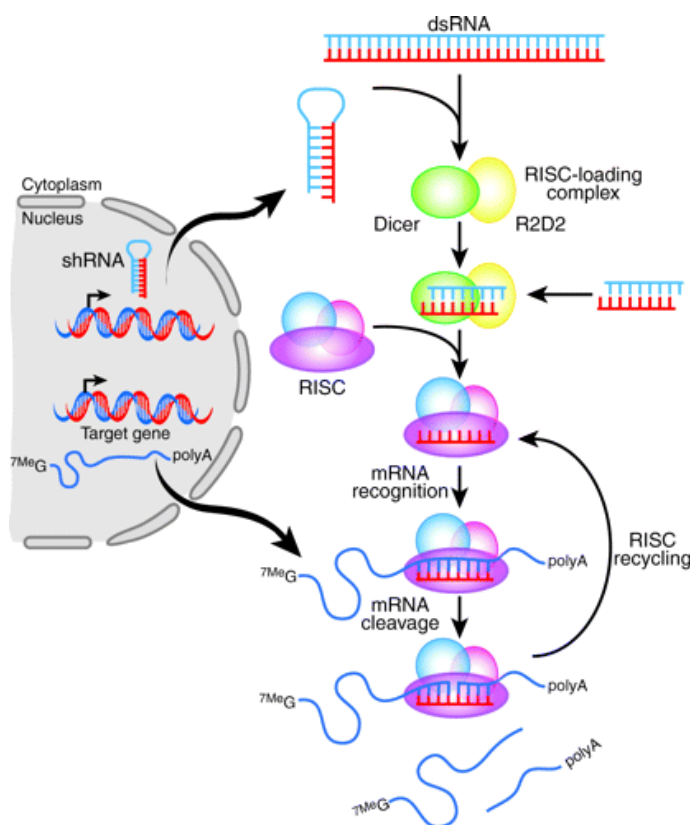
Dans le silencing siARN, l'ARN double brin est le plus souvent exogène (virus à ARN) ou résulte de l'action d'un expérimentateur comme dans le cas des expériences réalisées par Fire et Melo (ARN double brin de plusieurs centaines de nucléotides). Il fallait donc une étape pour expliquer l'apparition de petits ARN d'une vingtaine de nucléotides après l'introduction de ces longues séquences. Cette étape dans la biogénèse des petits ARN est réalisée par une enzyme de la famille des RNase III essentielle pour le développement, Dicer (l'éminceuse) (Bernstein *et al*, 2004). Dicer va générer à partir d'une matrice d'ARN double brin des ARN interférents d'une vingtaine de nucléotides.

L'ensemble de ces découvertes a permis de dégager un schéma général du mécanisme siARN (Fig. 1). L'idée est que l'ARN double brin de grande taille introduit dans la cellule (artificiellement ou issu de virus à ARN) est pris en charge par Dicer qui va l'hydrolyser en petits ARN d'une vingtaine de nucléotides. Ces derniers sont alors incorporés au complexe RISC en s'associant à la protéine Ago2. Dans le complexe RISC, seul le brin « guide » est conservé. Le brin « passager » issu du double brin formé par Dicer est lui éliminé soit par hydrolyse soit par dissociation (Matranga *et al*, 2005). Le choix du brin guide est déterminé



# Introduction

par plusieurs critères dont la stabilité de l'appariement des extrémités du petit ARN (Noland *et al*, 2013). C'est à ce moment que le complexe RISC est fonctionnel et peut interagir avec l'ARN messenger cible. C'est la complémentarité parfaite entre la séquence de l'ARN messenger cible et le petit ARN « guide » qui va permettre la coupure endonucléolytique par la protéine Argonaute. La coupure va permettre la dégradation des deux fragments d'ARN générés, par l'action des exonucléases 3'→5' et 5'→3' ce qui conduit à une disparition complète de l'ARN cible.



**Figure 1 : Régulation par les siARN.**

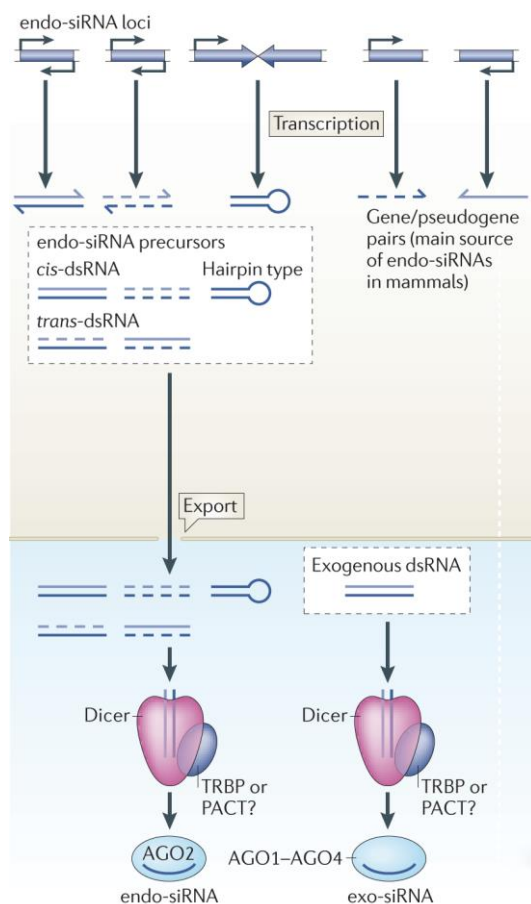
(Dominska et Dyckhoorn, 2010) Pour initier l'interférence ARN le précurseur ARN double brin est introduit dans la cellule soit directement soit par l'intermédiaire d'un gène exprimant un shRNA préalablement introduit dans la cellule. Ce précurseur est pris en charge par l'enzyme DICER (DICER/R2D2 chez *Drosophila*) qui va alors le découper en fragments de 21 paires de bases (siARN) qui sont par la suite incorporés dans le complexe effecteur RISC. Le siARN va alors servir de guide au complexe RISC qui va pouvoir se positionner sur l'ARNm du gène cible par complémentarité de bases. Une fois positionné, le complexe RISC via la protéine Argonaute 2 va alors couper l'ARNm cible le rendant sensible à l'action des exonucléases et empêchant sa traduction.

Des ARN double brin d'origine endogène peuvent aussi déclencher le mécanisme siARN via les endo-siARN. Chez les mammifères, du fait de l'absence d'ARN Polymérase dépendant de l'ARN, la formation d'ARN double brin est possible dans trois contextes génomiques (Fig. 2). (i) La transcription convergente sur un locus (cis-NAT) (ii) L'interaction entre deux séquences complémentaires transcrites par différents loci comme des séquences d'éléments transposables ou des pseudogènes (trans-NAT) (iii) La formation d'une structure ARN en tige boucle, distincte des précurseurs miARN, pouvant dériver de la répétition inversée d'une séquence. Dans chacun de ces cas, le précurseur ARN double brin va entrer dans le



# Introduction

mécanisme « classique » des siARN par action de la protéine Dicer et permettre la régulation des gènes cibles (Tam *et al*, 2008 ; Watanabe *et al*, 2006). Le rôle des endo-siARN n'est pas restreint à une action antivirale. Une étude portant sur les oocytes de souris a montré que l'absence des protéines Dicer et Ago2 résulte en la dérégulation des éléments transposables (Watanabe *et al*, 2008). Les endo-siARN auraient ainsi un rôle redondant avec celui des piARN dans la protection de la lignée germinale contre les éléments transposables. De la même façon que pour la voie des piARN, ce mécanisme semble être tissu-spécifique et restreint aux cellules de la lignée germinale (Faghihi *et al*, 2006 ; Werner *et al*, 2014).



**Figure 2 : Origines des endo-siARN** (Siomi *et al*, 2011). Les origines possibles des siARN au sein de la cellule. Ils peuvent être issus de transcrits convergents recouvrant l'ensemble ou des parties d'un gène. Ils peuvent également être produits à partir de transcrits formant des tiges boucles comme c'est le cas des précurseurs microARN. Enfin ils peuvent être issus de transcription inverse se produisant sur un pseudogène formant alors un couple sens/antisens. Dans tous ces cas le précurseur est pris en charge par Dicer qui le découpe en fragment de 21 paires de bases pour ensuite l'incorporer au complexe RISC.

## 1.3 - L'interférence par les miARN

La régulation par les microARN ou miARN va se différencier du silencing siARN par le mécanisme mis en œuvre, mais aussi par la biogénèse du petit ARN. En effet, contrairement





# Introduction

---

au mécanisme siARN qui est déclenché par l'introduction d'ARN double brin bimoléculaire et principalement d'origine exogène, le silencing miARN implique une séquence endogène simple brin transcrite à partir de gènes cellulaires (voir cependant l'exception des endo-siARN issus d'une longue tige boucle).

Historiquement, la première description d'un miARN a été faite par le groupe d'Ambros qui a mis en évidence que le gène *lin-4*, impliqué dans le contrôle du développement larvaire chez *C. elegans*, ne code aucune protéine mais permet l'accumulation d'un petit ARN (Lee *et al*, 1993). La génétique ayant mis en évidence que *lin-4* régulait négativement l'expression de plusieurs autres gènes et ce via la région 3' non traduite de leur ARNm, la connaissance de la séquence de *lin-4* a permis d'identifier la présence de plusieurs séquences partiellement complémentaires dans la région 3' non traduite des gènes cibles comme, *lin-14* (Lee *et al*, Wightman *et al*, 1993). Dans ces travaux, l'expression du petit ARN *lin-4* a un impact sur la quantité de protéines LIN-14 sans modifier pour autant la quantité d'ARN messager. Le modèle proposé est donc que le petit ARN transcrit par le gène *lin-4* se fixe en région 3'UTR de l'ARN messager du gène *lin-14* et bloque sa traduction.

L'ARN *lin-4* est maintenant devenu le prototype des miARN (Lagos-Quintana *et al*, 2001). C'est aussi le cas du gène *let-7*, autre gène impliqué dans le développement de *C. elegans* qui permet lui aussi l'expression d'un petit ARN de 22 nucléotides (pour la forme la plus abondante) et qui possède le même mode d'action que le miARN *lin-4* (Reinhart *et al*, 2000). Des homologues du gène *let-7* ont été identifiés chez l'homme, la drosophile et d'autres animaux (Pasquinelli *et al*, 2000). Cette classe très abondante de petits ARN est impliquée dans la régulation de nombreuses voies cellulaires (différenciation, prolifération, apoptose, etc...) et cela chez de nombreux organismes. Il existe ainsi près de 2000 gènes codant des miARN chez l'homme produisant 2600 micro ARN matures distincts tandis que chez *C. elegans* et la drosophile il existe 434 et 433 miARN, respectivement (miRBase). Ces gènes « miARN » ne codent pas directement des petits ARN mais leurs précurseurs aussi appelés pri-miARN (Lee *et al*, 2002). Ce transcrit primaire est dans la plupart des cas un produit de transcription de l'ARN Polymérase II et est coiffé et polyadénylé (Saini *et al*, 2005). Il existe néanmoins des cas particuliers où les pri-miARN sont transcrits par l'ARN Polymérase III, impliqués habituellement dans la transcription de l'ARN ribosomal 5 S, des ARN de transfert mais également d'autres petits ARN (snoARN entre autres) (Barchert *et al*, 2006). On





# Introduction

---

retrouve également des séquences de pri-miARN dans des séquences de gènes classiques. C'est alors l'ARN messager du gène dans lequel se trouve la séquence qui va avoir le rôle de précurseur. Ce cas de figure de miARN « intronique » (nom trompeur car la séquence du précurseur miARN est parfois dans le dernier exon) représente en moyenne 50% des miARN (Saini *et al*, 2007). Dans le cas de miARN « introniques », le niveau d'expression est souvent corrélé à celui de l'ARN messager hôte, signe que l'ARN messager joue bien dans ce cas-là le rôle de pri-miARN. Néanmoins, il existe des cas de promoteurs internes aux introns qui permettent une transcription de pri-miARN indépendante de l'ARN messager (Monteys *et al*, 2010).

Indépendamment de son origine, le pri-miARN est un transcrit très structuré qui peut contenir une ou plusieurs structures en tige-boucle d'environ 70 nucléotides. Chacune de ces tiges-boucles va être à l'origine d'un miARN. Les miARN sont souvent organisés en clusters et ainsi un pri-miARN va pouvoir générer jusqu'à 6 miARN différents (Alturia *et al*, 2005). La première étape dans la biogénèse du miARN va être la coupure des tiges-boucles du pri-miARN à leur base par le « micro-processeur ». Le complexe micro-processeur est composé de Drosha qui est une endonucléase spécifique à l'ARN double brin de la famille des RNase III (l'autre membre de la famille chez les mammifères avec Dicer) et DGCR8 (ou Pasha chez la drosophile) protéine indispensable à la stabilisation de Drosha sur son substrat (Denli *et al*, 2004 ; Gregory *et al*, 2004). Ce complexe permet la coupure à 11 nucléotides de la base de la tige-boucle donnant naissance au précurseur intermédiaire pré-miARN. L'action du complexe micro-processeur se fait dans le noyau et les pré-miARN ainsi produits sont pris en charge par une voie de transport vers le cytoplasme associée à l'exportine 5 (Lund *et al*, 2004 ; Yi *et al*, 2003). L'exportine 5 est une protéine de la famille de l'importine  $\beta$  qui permet le transport vers le cytoplasme via l'utilisation de RanGTP comme énergie. Elle reconnaît les pré-miARN grâce à la signature de coupure par Drosha qui génère une extrémité 3'OH protubérante de deux nucléotides (Okada *et al*, 2009).

Une fois dans le cytoplasme, les pré-miARN sont pris en charge par Dicer. Dicer va agir sur le pré-miARN comme sur un ARN double brin dans la biogénèse des siARN en produisant des duplex miARN d'une vingtaine de nucléotides. Chez les mammifères, il n'existe qu'une seule protéine Dicer mais chez la Drosophile il existe deux paralogues qui sont spécialisés : Dicer 1 est spécifique de la coupure des pré-miARN et va permettre le transfert préférentiel du petit



# Introduction

---

ARN produit dans Ago1 et Dicer 2, spécifique de l'ARN double brin et du silencing siARN, va permettre un transfert préférentiel du petit ARN dans Ago2. Chez les mammifères, il n'y a pas d'indication que les duplex miARN soient intégrés préférentiellement à une protéine Argonaute particulière. Néanmoins, le modèle proposé est que seule Ago2 est capable d'éliminer le brin « passager » par hydrolyse. Les autres protéines Argonaute dépendent d'une déstabilisation du duplex due à la présence de mésappariement. Les duplex miARN contenant quelques mésappariement, on peut s'attendre à ce que les complexes RISC actifs formés avec les protéines AGO 1, 3 et 4 soient plus enrichis en miARN. A l'inverse, même si cela n'a pas été démontré, il est supposé que les endo-siARN sont préférentiellement transférés dans Ago2.

La compréhension du mode d'action des miARN a été fortement influencée par les premières études faites sur la régulation de *Lin-14* par *Lin-4* (Banerjee et Slack, 2002). Ces études montrent que le miARN *lin-4* est capable de former des hybrides incomplets avec l'ARN messager de *lin-14* dans sa région 3'UTR non traduite (7 sites de fixation). La présence de plusieurs sites de fixation laisse à penser qu'il est nécessaire de recruter plusieurs complexes RISC sur le transcrit pour permettre une réponse efficace. La particularité des miARN est que l'hybridation avec leur cible est imparfaite, du moins chez les mammifères, où il n'existe que très peu de cas d'interaction avec une cible parfaite (Yekta *et al*, 2004), à l'inverse des plantes. En conséquence, le nombre de nucléotides impliqués dans la reconnaissance de la cible des microARN est plus faible que dans le cas des siARN et un miARN peut posséder plusieurs ARN messagers cibles. Typiquement on peut s'attendre à ce que le motif clé de la reconnaissance (la séquence d'amorçage, « seed sequence », les nucléotides 2 à 8 du micro ARN) soit présent tous les 16 kb de séquence ou environ dans un ARN messager sur 10. Une contrepartie de cette possibilité d'une multiplicité de cible est que l'absence d'une complémentarité parfaite avec sa cible ne permet pas l'hydrolyse de l'ARN messager cible mais fait intervenir d'autres mécanismes. Néanmoins, chez les mammifères, il n'existe pas de séparation stricte entre les siARN et les miARN et il existe des cas, comme celui du miR-196a qui cible HoxB8 et qui ne diffère que d'une paire de base avec son ARN messager cible, où le mécanisme impliqué est une hydrolyse de l'ARN (Yekta *et al*, 2004). Enfin si la majorité des sites miARN est retrouvée dans la région 3' non traduite des ARNm, des sites peuvent néanmoins être présents dans la région 5' non traduite ou dans



les régions codantes (Brodersen et Voinnet 2009). Ainsi il convient d'être prudent avant de trop généraliser les règles principalement issues des études faites sur *C.elegans* et les différences observées dans les études de la régulation par le complexe miRISC ont longtemps généré des controverses. Les premières observations avait montré par exemple que chez *C. elegans* les transcrits *lin-14* n'étaient pas dégradés et restaient associés à des polysomes (Wightman *et al*, 1993). Ceci suggérait une régulation de la traduction au niveau de l'élongation ou une dégradation co-traductionnelle de la protéine (Olsen et Ambros, 1999). Mais d'autres études, y compris chez *C. elegans*, ont montré une dégradation de l'ARN messenger (Bagga *et al*, 2005) ou l'inhibition de l'initiation de la traduction.

Aujourd'hui plusieurs milliers d'études chez *C. elegans*, la drosophile et les vertébrés ont permis d'éclaircir le tableau. Ainsi, la présence de sites de fixation miARN sur l'ARN messenger cible peut conduire à une diminution de l'expression protéique qui est le plus souvent associée à une diminution de l'ARN messenger (Eichhorn *et al*, 2014 ; Guo *et al*, 2010). Cependant contrairement aux mécanismes impliqués dans le silencing par un siARN, pour le silencing miARN il n'y a pas d'indication d'hydrolyse endonucléolytique de l'ARN messenger cible mais plutôt une dégradation par la voie normale de dégradation après déadénylation et décoiffage de l'ARN (Wu *et al*, 2006) (Fig. 3). L'inhibition de la traduction ne permet pas à elle seule d'induire une déadénylation du transcrit. Il est donc nécessaire au complexe miRISC de recruter d'autres protéines partenaires. La déstabilisation de l'ARN messenger par le miARN se fait via le recrutement dans le complexe miRISC de CCR4/NOT (chez la drosophile, Behm-Ansmant, 2006 ; chez les mammifères, Chen *et al*, 2009). CCR4/NOT est le complexe protéique qui joue un rôle déterminant dans la déadénylation processive des ARN messagers et donc leur dégradation « active ». Cette déadénylation du message par le complexe miRISC va s'accompagner du recrutement du complexe DCP1/DCP2. Ce complexe va permettre l'élimination de la coiffe en 5' de l'ARN messenger. La combinaison de l'action de ces deux complexes va rendre alors le transcrit ciblé par le miARN sensible aux exonucléases cellulaires et permettre sa dégradation (Chen *et al*, 2009) (Fig. 3).

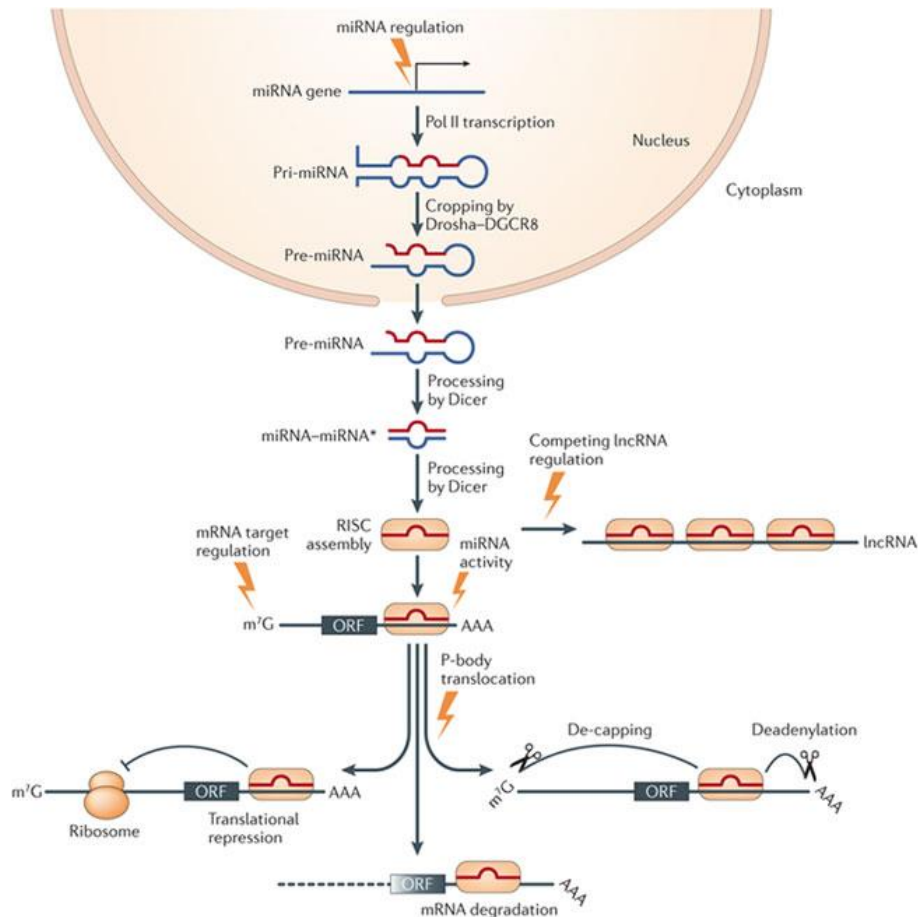
Contrairement au silencing siARN, ou le complexe RISC minimal est uniquement composé de la protéine Ago2 (porteuse de l'activité endonucléolytique) et du petit ARN, le complexe miRISC nécessite plusieurs partenaires comme CCR4/NOT et DCP1/DCP2. Le recrutement de ces complexes protéiques est donc primordial dans la mise en place du silencing miARN et



passer par les protéines de la famille GW182 (TNRC6 A, B et C chez l'homme). Cette famille de protéines est capable d'interagir avec les protéines Argonaute via un motif répété riche en Glycine-Tryptophane (GW) auquel elles doivent leurs noms (Till *et al*, 2007). La protéine GW182 ou TNRC6 va avoir un rôle de plateforme permettant le recrutement de tous les facteurs nécessaires au silencing. Des expériences d'ancrage de la protéine GW182 chez la drosophile (Behm-Ansmant *et al*, 2006) et de mammifère (Li *et al*, 2008) ont montré l'importance de cette protéine, car cet ancrage permet l'induction d'un silencing semblable à celui d'un miARN mais indépendamment des protéines Argonaute et du miARN (Lazaretti *et al*, 2009). Ceci indique que GW182 suffit à apporter la fonction de silencing au sein du complexe miRISC.

Les protéines GW182 sont présentes dans le cytoplasme à la fois sous forme soluble mais aussi dans des agrégats de protéines et d'acides nucléiques appelés P-bodies (Processing Bodies, selon la nomenclature initialement introduite chez la levure (Sheth *et al*, 2003)). Ces structures contiennent les enzymes de la machinerie de dégradation des ARN comme Xrn1 (l'exonucléase 5' présente dans le cytoplasme) ou DCP1 (l'enzyme de décoiffage des ARNm) (Eystathiou *et al*, 2003). La présence des protéines Argonaute a aussi été mise en évidence dans les P-bodies ce qui pourrait suggérer un rôle de ces structures dans la régulation par les petits ARN (Liu *et al*, 2005 ; Sen et Bau, 2005). Cependant, la présence dans les cellules de mammifère de P-bodies détectables ne semble pas nécessaire pour le silencing par les siARN et les miARN (Serman *et al*, 2007) ce qui suggère plutôt un rôle dans le stockage, le recyclage ou l'assemblage des constituants de la machinerie de silencing (Leung et Sharp, 2013).





Nature Reviews | Drug Discovery

**Figure 3 : Biogenèse et mécanisme d'action des miARN** (Van Rooij et Olsen, 2012).

Les gènes des miARN sont généralement transcrits par l'ARN polymérase II qui produit un transcrit coiffé et polyadénylé, le pri-miARN. Ce pri-miARN possède des structures en tige-boucle qui vont être clivées à leur base par la RNase III Drosha associée à DGCR8 générant les pre-miARN, précurseurs de 70 nucléotides en tige-boucle. Après avoir été exporté dans le cytoplasme, le pre-miARN est coupé par Dicer pour donner un miARN double brin d'environ 22 nucléotides. Un de ces deux brins miARN est incorporé au complexe RISC et permet de guider le complexe effecteur sur l'ARNm cible. Le complexe miRISC va modifier le devenir de l'ARNm de plusieurs manières à travers une inhibition de la traduction de l'ARNm cible mais également sa déadénylation et son décoiffage conduisant à sa dégradation par les exonucléases de la cellule.



## 1.4 – L'interférence par les piARN

Les « PIWI-interacting small RNAs » ou piARN sont la classe la moins bien caractérisée des petits ARN régulateurs liés à l'interférence ARN. Ils ont pourtant un rôle crucial de « gardien du génome » en maintenant une répression de l'ensemble des éléments transposables. Le manque de connaissance sur ce mécanisme provient sans doute du fait que ce mécanisme est restreint aux cellules de lignée germinale, même si quelques travaux suggèrent qu'ils pourraient avoir des fonctions non germinales chez l'Aplysie (Rajasethupathy *et al*, 2012), la drosophile (Rouget *et al*, 2010) ou les mammifères (Watanabe *et al*, 2011). En dépit de cette spécialisation, cette voie de régulation a été observée très tôt dans l'histoire de l'interférence ARN. Dès 2001, une étude de la lignée germinale mâle de *D. melanogaster* montre comment la protéine *stellate*, qui a la particularité d'être codée à partir de séquences répétées, est mis sous silence par un petit ARN (Aravin *et al*, 2001). Il a d'ailleurs été observé très vite que dans les organes sexuels et le jeune embryon de *D. melanogaster* a été trouvée une forte abondance de ces petits ARN d'une taille plus importante que celle des siARN et miARN (23-32 nucléotides). Ces petits ARN étant dérivés de séquences d'éléments répétés dans le génome, ils furent alors nommés dans un premier temps « repeat-associated small interfering RNAs » ou rasiARN (Aravin *et al*, 2003).

Comme c'est chez la drosophile qu'il existe le plus de données sur les piARN, le reste de cette présentation considère principalement le cas de la drosophile. Les piARN sont codées à partir de longues séquences intergéniques dans le génome (> 100 000 bases) aussi appelées piARN clusters et souvent très riches en éléments transposables. La majorité des piARN sont issus du brin antisens des séquences des éléments transposables permettant un silencing via l'hybridation avec le transcrit des transposons (Saito et Siomi, 2010).

Contrairement à la biogénèse des siARN et miARN, le précurseur piARN est un ARN simple brin dont la maturation va être indépendante d'enzymes de la famille des RNases III de type Dicer ou Drosha (Vagin *et al*, 2006). L'autre différence majeure avec les siARN et miARN est que les piARN ne sont pas associés aux protéines de la sous-famille Argonaute qui sont exprimées de manière ubiquitaire mais aux protéines de la sous-famille PIWI. La sous-famille PIWI est principalement exprimée dans la lignée germinale. L'interaction des piARN avec les protéines PIWI permet la formation du complexe piRISC qui participe à deux mécanismes distincts. Un mécanisme primaire qui va permettre le silencing des éléments transposables



# Introduction

---

au niveau transcriptionnel et un second qui va permettre l'amplification de la réponse contre ces éléments via la génération de nouveaux piARN à partir des transcrits cibles.

Le transcrit précurseur piARN ou pre-piARN est transcrit à partir d'un piARN cluster (Fig. 4). Cet ARN est exporté et maturé dans le cytoplasme. La première étape de la maturation sera être coupée pour donner des « piRNA like small RNA » qui sont les piARN non encore disponibles pour le complexe RISC. L'activité endonucléolytique est apportée par la protéine mitochondriale transmembranaire Zucchini ou Zuc. On ignore encore la séquence définissant le site de clivage au sein des pre-piARN mais il est clair qu'il existe un très fort biais d'incorporation dans les protéines PIWI de piARN possédant un 5' Uracile. Une des signatures de la protéine Zuc va être d'ailleurs de créer cette extrémité 5' monophosphate Uracile qui va permettre l'incorporation dans la protéine PIWI (Ipsaro *et al*, 2012 ; Nishimasu *et al*, 2012). Une fois incorporé dans le complexe piRISC le piARN subi encore deux étapes de maturation. La première sera un raccourcissement de son extrémité 3' par l'action d'une exonucléase qui réduit sa taille à 23-32 nucléotides (Kawaoka *et al*, 2011). Des travaux récent ont montrés que cette exonucléase été également la protéine Zuc et que sa coupure libérais de nouveaux piARN primaires à l'extrémité 5' monophosphate Uracile aussi appelés Trail-piARN qui permettent une amplification de la réponse (Mohn *et al*, 2015). La seconde sera la 2'-O méthylation de cette même extrémité 3' du piARN afin de protéger cette extrémité de la dégradation et de stabiliser le complexe piRISC. Cette étape de la maturation sera permise par l'action des protéines Hen1/Pimet chez *D. melanogaster* (Horwich *et al*, 2007). C'est à la fin de cette dernière étape de maturation que le complexe piRISC devient actif.

Le complexe piRISC est composé de nombreuses protéines dont le rôle dans la régulation par les piARN n'a pas encore été déterminé pour toutes. La particularité des protéines PIWI est que ces dernières, contrairement aux protéines Argonaute, présentent des méthylations qui permettent le recrutement d'une classe de protéines à domaine Tudor (TDRD) spécifiques elles aussi de la lignée germinale (Siomi *et al*, 2010). C'est le cas de *Tud* qui s'associe avec Aubergine via ses méthylations pour participer aux mécanismes d'amplification de la réponse piARN (Nishida *et al*, 2009) chez la drosophile. La mutation de *Tud* engendre le même phénotype que la mutation d'Aubergine (Arkov *et al*, 2006).

Le mécanisme impliqué par le complexe piRISC dépend de la protéine PIWI associée. Si le



# Introduction

---

complexe piRISC est composé autour de la protéine PIWI cela déclenche le mécanisme primaire piARN, le complexe sera relocalisé dans le noyau (Haase *et al*, 2010 ; Olivieri *et al*, 2010 ; Saito *et al*, 2010) et permettra le recrutement de méthyltransférases qui induiront le silencing des promoteurs des éléments transposables (Aravin *et al*, 2008 ; Le Thomas *et al*, 2013). Dans le cas de la formation du complexe RISC associé à la protéine Aubergine cela engendre le mécanisme d'amplification (Fig. 4). Ce mécanisme est cytoplasmique et vise les ARN issus des éléments transposables. Les ARN ciblés par le complexe piRISC seront clivés via une action endonucléolytique (Saito *et al*, 2006). Le clivage s'effectue toujours au niveau du 10<sup>e</sup> nucléotide qui s'apparie avec le piARN à partir de l'extrémité 5' (Brennecke *et al*, 2007 ; Gunawardane *et al*, 2007). Le clivage de l'ARN du transposon cible crée une extrémité 5' sur celui-ci donnant ainsi le précurseur d'un nouveau piARN. Ce dernier sera alors pris en charge par la protéine Ago3 et sera maturé au niveau de son extrémité 3' comme pour un piARN primaire même si l'intervention des mêmes partenaires n'est pas claire. Le nouveau complexe piRISC associé avec Ago3 agit alors à son tour et de la même façon qu'avec Aubergine sur l'ARN du transposon cible donnant naissance à un piARN secondaire qui sera pris en charge par Aubergine permettant une amplification cyclique (Brennecke *et al*, 2007 ; Gunawardane *et al*, 2007). C'est ce renvoi mutuel de substrat entre Aubergine et Ago3 qui a inspiré le nom de mécanisme Ping-Pong. La mise en place de ce mécanisme d'amplification est facilité par la protéine à domaine Tudor, Krimper, qui permet la colocalisation des protéines Ago3 et Aubergine sein de même complexe (Webster *et al*, 2015)

A l'image des P-bodies dans le silencing miARN, la machinerie piARN se localise dans le cytoplasme dans des granules dépourvus de membranes. Ces structures non membranaires riches en ARN sont conservées dans les lignées germinales de nombreuses espèces d'eucaryotes. Chez *D. melanogaster*, où ces granules ont été le plus étudiés, il est observé qu'ils sont regroupés en une grande structure périnucléaire dense aux électrons appelée « nuage » (Findley *et al*, 2003). Nuage est une structure uniquement retrouvée dans les cellules nourricières de l'oocyte et l'ensemble de la lignée germinale mais absente de l'oocyte (cellule quiescente) et des cellules du soma. L'ensemble des facteurs nécessaires au silencing piARN telles que les protéines PIWI et les protéines à domaine Tudor s'accumule dans cette structure. Nuage colocalise avec des mitochondries permettant la présence de la



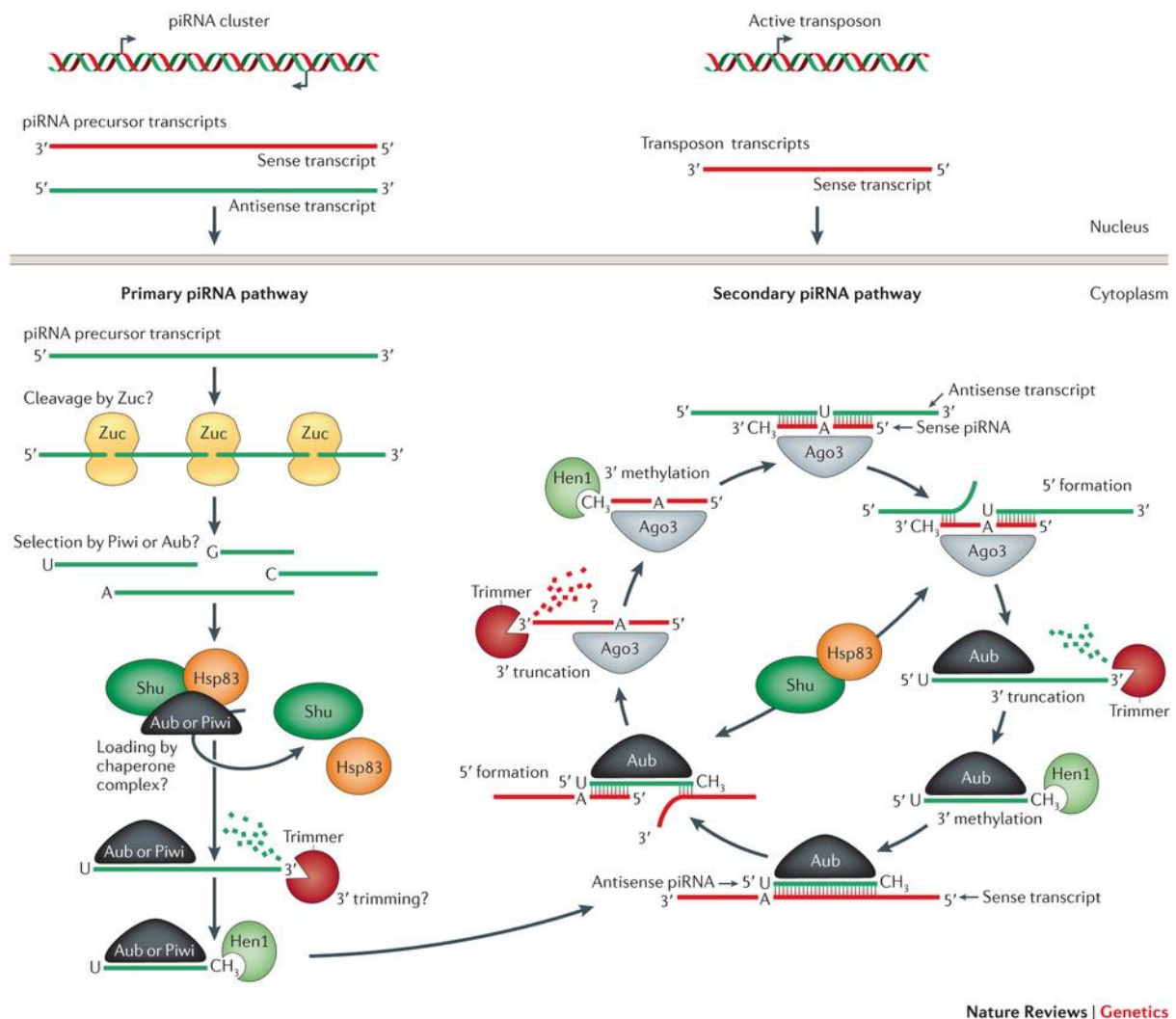


protéine transmembranaire Zuc, essentielle pour la biogénèse du piARN. Chez les souris, on retrouve le même type de granules où s'accumulent les constituants de la machinerie piARN. Ils portent le nom de « ciment intermitochondrial » ou « chromatoid bodies (CBs) » (Kuramochi-Miyagawa *et al*, 2001 ; Vasileva *et al*, 2009). Que cela soit chez la drosophile ou la souris, la structure et les composants de ces granules vont être modifiés au cours du temps selon les différents stades de différenciation des cellules de la lignée germinale.

Les protéines de la sous-famille Piwi ont une activité endonucléase et peuvent donc *a priori* être impliquées dans un silencing post-transcriptionnel, c'est par exemple ce qui est observé pour le silencing de *Stellate* chez la drosophile (Kotelnikov *et al*, 2009). Dans ce cas, il est proposé que seuls les ARN compétents pour la traduction soient dégradés par le complexe Aub/piARN ce qui conduit à un silencing ARN nettement plus faible que celui observé pour les protéines. Chez la souris, une activité de coupure dirigée par les piARN a été observée au cours de la spermatogenèse (Goh *et al*, 2015). Par contre, d'autres silencings ont lieu au niveau transcriptionnel comme évoqué plus haut ou dans le cas du locus *Rasgrf1* qui est soumis à empreinte parentale chez les mammifères (Watanabe *et al*, 2011).



# Introduction



Nature Reviews | Genetics

**Figure 4 : Biogenèse primaire et secondaire des piARN** (Luteijn et Ketting, 2013).

Exemple des deux voies de biogenèse des piARN chez *Drosophila melanogaster*. Dans la voie de synthèse primaire, le précurseur issu des gènes piARN est clivé par l'endonuclease Zucchini (Zuc). Aidés par deux protéines chaperonnes Shutdown (Shu) et Heat shock protein 83 (Hsp83), les fragments sont incorporés dans les protéines Piwi ou Aubergine. De préférence les fragments aux extrémités 5'U sont incorporés. Une fois incorporé, le précurseur piARN continue d'être maturé en subissant un raccourcissement de son extrémité 3' par une exonucléase inconnue puis une méthylation de cette même extrémité par l'enzyme Hen1. Cela clôt la biogenèse primaire. Si le piARN est associé avec la protéine Aubergine (mais pas Piwi) celui-ci peut participer à la synthèse secondaire. En reconnaissant le transcrit cible, Aubergine permet sa coupure. L'extrémité 5' libérée sur le transcrit cible va alors servir de nouveau précurseur et sera incorporée à la protéine Ago3. Ago3 permet le même mécanisme de coupure donnant à son tour des précurseurs pour de nouveaux piARNs à Aubergine permettant l'amplification de la réponse.



## 2 – Les longs ARN non codants (lncRNA)

### 2.1 – Première description et caractéristiques

L'essentiel du génome ne code aucune protéine et la découverte des petits ARN et de leurs précurseurs ne peut suffire à attribuer une fonction à ces régions sombres du génome, surnommées la « Dark matter » (Johnson *et al*, 2005). Lors de la dernière décennie, l'avancée de la technologie de séquençage a grandement aidé à éclairer cette part d'ombre de l'ADN. Ceci a mis en lumière qu'en plus du transcriptome codant il existe un transcriptome non codant qui représente la presque totalité des séquences du génome. Si ces observations ont maintenant été largement validées expérimentalement, elles ne fournissent pas pour autant un cadre conceptuel pour interpréter le rôle biologique de l'ensemble du génome. De fait, pour de nombreux transcrits, leur faible niveau d'accumulation est compatible avec la proposition qu'ils n'ont pas d'activité biologique mais traduisent simplement les limites de la fidélité de l'initiation de la transcription (Struhl *et al*, 2007) ; une hypothèse renforcée par le très petit nombre de mutations dans des ARN non codants qui ont été associées à des pathologies (Ponting *et al*, 2009). Alternativement, comme certains de ces transcrits sont rapidement dégradés, il est possible que ce soit la transcription plus que les transcrits produits qui joue un rôle biologique. C'est pourquoi, si un répertoire exhaustif des transcrits non codants devrait pouvoir être établi dans les années qui viennent, rien ne permet aujourd'hui de prédire le nombre de classes fonctionnelles qu'il contiendra, ni à quel pourcentage des ARN non codants on pourra *in fine* associer une activité biologique.

Pour distinguer ces transcrits des petits ARN précédemment discutés nous utiliserons l'appellation générique de longs ARN non codants ou lncRNA (long non coding RNA). La première description d'un long ARN non codant ayant une activité de régulation de l'expression génique est celle de *Xist* (Bursani *et al*, 1991 ; Brown *et al*, 1991 ; Brockdorff *et al*, 1992). Ce dernier est impliqué dans l'inactivation du chromosome X surnuméraire chez les femelles des mammifères. *Xist* est un ARN de 17 kb exprimé à partir du locus XIC (X-inactivation center) du chromosome X inactif. L'expression de *Xist* est suffisante à elle seule pour permettre l'induction en *cis* d'un silencing complet du chromosome X (Wutz *et al*, 2000). Cette première description a permis d'identifier quelques caractéristiques des longs ARN non codants qui ont été reprises dans les travaux plus récents. De façon générale, les



# Introduction

---

longs ARN non codants sont transcrits par la polymérase II, sont coiffés et peuvent être polyadénylés et épissés. Par convention ils ont une taille supérieure à 200 nucléotides qui peut atteindre plusieurs centaines de kilobases. De plus, comme leur nom l'indique, ces transcrits ne codent pas de protéines, mais il est difficile de donner un sens absolu à cette définition puisqu'un ARN de grande taille contient nécessairement des cadres ouverts de lecture (Dinger *et al*, 2008). On est donc amené à qualifier le caractère non codant par une faible capacité de codage protéique : seuls des cadres de lecture courts sont détectables, qui ne sont pas conservés entre espèces et sont dans un environnement peu favorable à la traduction. Ceci est en général complété par l'absence de détection d'un produit protéique correspondant aux cadres ouverts les plus longs. Il est cependant souvent impossible de complètement exclure que des peptides pourraient être produits. Le développement de la technique d'empreinte ribosomique qui permet une description très fine de la traduction n'a pas complètement permis de répondre à la question de la non traduction des ARN non codants puisque l'on peut observer une association à bas niveau avec les ribosomes même pour des ARN « bien identifiés » comme non codants (Guttman *et al*, 2013). Le critère le mieux adapté pourrait être l'absence de diminution de l'association avec les ribosomes après un codon stop, ce qui indiquerait que l'association avec les ribosomes n'est pas liée à une synthèse protéique spécifique.

L'observation que les ARN non codants sont faiblement conservés entre espèces, en moyenne un peu moins que les régions non traduites des ARNm (Carninci *et al*, 2005) contribue à jeter un doute sur la fonction de ces transcrits. On peut cependant noter que les régions non codantes des ARNm remplissent des rôles importants pour l'expression des gènes à travers le contrôle de la stabilité, de la localisation et de la traduction de ces ARN. Néanmoins, l'analyse de la conservation des séquences ne permet que difficilement d'identifier les régions fonctionnelles du fait de la petite taille des motifs impliqués et de l'importance des structures secondaires mais pas nécessairement des séquences elles-mêmes.

Il a été proposé trois mécanismes pour la genèse de nouveaux ARN non codants (Ponting *et al*, 2009) :

(i) Un décalage du cadre de lecture dans un gène codant une protéine. C'est l'hypothèse qui est faite pour l'apparition de l'ARN *Xist*. Celui-ci semble dériver du gène codant *Lnx3* dont la



séquence a subi un décalage du cadre de lecture après insertion d'un élément transposable tôt dans l'histoire évolutive des mammifères placentaires (Duret *et al*, 2006 ; Elisaphenko *et al*, 2008).

(ii) La seconde possibilité résulte de réarrangements chromosomiques pouvant permettre la transcription d'ARN non codant par l'apport d'un promoteur dans des régions non transcrites précédemment (Ponting *et al*, 2009 ; ESTs BM537447, C0597044, et DN744681)

(iii) Il est possible également que certains longs ARN non codants soit apparus suite à des événements de rétrotranspositions ; l'apparition d'un pseudogène non codant qui au cours du temps acquiert de nouvelles fonctions. Ainsi, l'insertion d'un promoteur sur le brin opposé d'un gène peut donner naissance à un transcrit non codant antisens (Conley *et al*, 2008).

Quelles que soient leurs origines, on retrouve des longs ARN non codants dans l'ensemble des tissus, certains étant exprimés de façon tissu spécifique (Mercer *et al*, 2008). Leur localisation dans la cellule va aussi varier et, probablement, reflète leur fonction. Un exemple bien étudié est « Gomafu » ou RNCR2 (Retinal Non Coding RNA 2). Cet ARN non codant s'exprime uniquement dans les cellules du système nerveux (Ishii *et al*, 2006) et s'accumule dans le noyau au niveau des « nuclear speckles » (Sone *et al*, 2007). Cette localisation bien spécifique correspond à sa fonction proposée de régulateur de l'épissage (Tsuiji *et al*, 2011).

Compte tenu de la diversité potentielle des longs ARN non codants, de leurs fonctions, de leurs mécanismes d'action et de leurs interactions avec leurs cibles, plusieurs classifications des longs ARN non codants ont été proposées, le plus souvent en fonction de leurs positionnements par rapport aux gènes codants déjà identifiés dans le génome, ce qui conduit à quatre catégories principales :

(i) La catégorie « sens », le long ARN non codant recouvre l'exon d'un autre gène sur le même brin (Kapranov *et al*, 2005).

(ii) La catégorie « antisens » (asRNA ou Natural antisense transcript, NAT), le long ARN non codant recouvre l'exon d'un autre gène mais sur le brin opposé (Katayama *et al*, 2005). Cela permet en théorie une hybridation entre l'ARN non codant et le transcrit « sens » du gène recouvert. Une partie plus complète sur cette catégorie spécifique d'ARN, qui est le sujet de ma thèse, sera commentée plus tard dans cette présentation.



(iii) La catégorie « intronique » (sisRNA, stable intronic sequence RNA ; TIN et PIN, totally or partially intronic RNA), le long ARN non codant dérive de l'intron d'un autre transcrit. Son expression va donc dépendre du gène dont il est l'hôte (Nakaya *et al*, 2007 ; St Laurent *et al*, 2012 ; Gardner *et al*, 2012).

(iv) La catégorie « intergénique » (lincRNA, long intergenic non coding RNA), le long ARN non codant est localisé entre deux gènes. Il possède son propre promoteur et est totalement indépendant (Guttman *et al*, 2009).

En plus de cette classification, le mode d'action de ces ARN va permettre de les classer au sein de chaque catégorie. Deux classes de mode d'action sont souvent utilisées : soit le long ARN non codant a un effet local en *cis* sur la région où il est transcrit, soit il a un effet à distance en *trans* sur un ou des gènes d'un autre locus.

## 2.2 – Une multiplicité de modes d'action

Pour distinguer les lncRNA « fonctionnels » du bruit de fond transcriptionnel il sera nécessaire d'étudier en détail chacun de ces transcrits. On peut cependant envisager que des critères permettent d'augmenter la probabilité qu'un lncRNA ait une fonction. A défaut d'une conservation de séquence entre espèces qui en général est très faible, la conservation d'une organisation génomique lorsqu'il existe des gènes « classiques » à proximité peut constituer une première indication. Le niveau d'expression est aussi un paramètre qui peut être utilisé, le séquençage à grande profondeur permettant d'identifier des transcrits dont l'abondance moyenne est inférieure à une molécule par cellule. Si de tels niveaux d'expression sont suffisants pour une caractérisation moléculaire du transcrit, ceci rend peu probable une activité biologique telles que plusieurs d'entre elles ont été identifiées jusqu'à présent. Il ne faut cependant pas exclure que le niveau d'expression puisse être plus élevé dans un type cellulaire particulier ou en réponse à des signaux spécifiques, et aussi que le mode d'action consiste en l'initiation d'une réponse qui se développe à travers un mécanisme d'amplification. En dehors des exemples historiques comme Xist et ceux qui seront décrits plus en détail ultérieurement, quelques études ont cherché à savoir quelle était la fraction des ARN non codants susceptibles d'avoir une activité biologique détectable. C'est en particulier un des objectifs des travaux du Broad Institute sur la caractérisation des lincRNA. Le choix de se concentrer sur les lincRNA reflète à la fois des aspects techniques et



biologiques. Par définition, les lincRNA sont situés loin des gènes classiques ce qui permet de diminuer la probabilité qu'il s'agisse d'un fragment d'un ARN messager qui ne contiendrait plus les régions codantes initialement présentes. De plus les lincRNA contiennent les marques chromatinienne d'un gène transcriptionnellement actif et sont exprimés à des niveaux un peu plus faibles mais néanmoins comparables à ceux des ARNm, l'organisation du gène étant conservée chez les mammifères. En se focalisant sur les lincRNA exprimés dans les cellules ES une étude de silencing systématique par siARN a montré que ceux-ci avaient un impact sur le transcriptome comparable à celui des facteurs de transcription y compris ceux connus pour contrôler l'identité des cellules souches (Guttman *et al*, 2011). Ainsi, pour peu que l'on ait sélectionné une bonne population de gènes candidats, les lincRNA peuvent avoir un impact biologique comparable à celui des gènes classiques.

Ces ARN agissent par des mécanismes d'actions très variés qui régulent l'expression des gènes à au moins trois niveaux différents : épigénétique, transcriptionnel et post-transcriptionnel. Afin de montrer toute cette diversité de régulation par les longs ARN non codants, une description de modèles représentatifs de chaque mode d'action va être présentée.

## 2.2.1 – Modulation de la chromatine

Une des caractéristiques de nombreux longs ARN non codants est leur capacité à recruter des complexes protéiques. Certains de ces complexes protéiques vont permettre d'agir sur les régulations épigénétiques (Khalil *et al*, 2009) (Tsai *et al*, 2010). Le recrutement des complexes modificateurs d'histones ou de méthylation de l'ADN est d'ailleurs l'un des mécanismes régulateurs fréquemment retrouvé chez les longs ARN non codants découverts à ce jour. C'est le cas des ARN Xist, HOTAIR, Kcnq1ot1 et Air qui sont connus pour interagir avec les complexes Polycomb PRC1 et PRC2 (Kaneko *et al*, 2010). Ces ARN peuvent agir aussi bien *en cis* qu'en *trans*.

### ***L'action en cis de l'ARN Xist***

L'ARN Xist est l'un des premiers exemples d'ARN permettant le recrutement d'un complexe modificateur d'histones. L'ARN Xist est un ARN non codant de 17 kilobases transcrit à partir



# Introduction

---

de « l'X-inactivation center » situé sur le chromosome X inactif. Cet ARN est essentiel au dosage et à l'inactivation du chromosome X surnuméraire chez les mammifères femelles. Les ARN Xist vont recouvrir le chromosome X à inactiver et induire son silencing. L'action de Xist se fait donc en *cis* - mais à l'échelle d'un chromosome entier - et conduit à l'inactivation du chromosome X qui l'exprime tout en épargnant l'autre chromosome X. Ce silencing est dû au recrutement par l'ARN du complexe modificateur d'histones PRC2 (Fig. 5). Ce recrutement se fait grâce à un motif répété situé à l'extrémité 5' de l'ARN connu sous le nom de répétitions A (repeats A) (Zhao *et al*, 2008). Le motif « repeats A » interagit directement avec la sous-unité catalytique de PRC2, EZH2 et permet la déposition sur le chromosome de la marque de l'hétérochromatine spécifique du complexe Polycomb qui est la triméthylation de l'histone H3 sur la lysine 27, H3K27me3 (Plath *et al*, 2003).

En plus de l'inactivation du chromosome X auquel il est associé, Xist semble capable de le relocaliser dans le noyau. Ceci est possible grâce au « nucleation center » que l'on retrouve dans l'exon 1 du gène codant Xist. Ce dernier permet le recrutement du facteur de transcription YY1 (Jeon *et al*, 2011). YY1 reconnaît 3 sites de liaison dans l'exon 1 (Fig. 5). Cette protéine a la particularité de pouvoir interagir aussi bien avec l'ADN qu'avec l'ARN et permet d'être l'adaptateur entre Xist et le chromosome X. Le chromosome X inactivé sera alors isolé de l'ensemble des autres chromosomes. L'adaptateur YY1 semble important dans l'inactivation du chromosome car la mutation de ces sites de fixation dans le « nucleation center » aura pour effet l'absence de silencing et étonnamment dans ce cas de figure particulier l'ARN Xist sera capable d'agir en *trans* sur l'autre chromosome X. Il est montré de plus que la protéine YY1 se fixe uniquement sur le « nucleation center » du chromosome X inactif et pas sur celui de l'autre. Il est probable qu'au moment de l'inactivation du chromosome X dans le développement, l'ensemble des sites de fixation d'YY1 soient bloqués mis à part ceux du chromosome X à inactiver (Jeon *et al*, 2011). Dans ce cas, Xist permet l'apport d'un complexe de modification des histones et son interaction avec un adaptateur protéique sur le chromosome permet d'expliquer la mise en place du silencing. Néanmoins, le mécanisme qui permet de recouvrir spécifiquement l'ensemble du chromosome X n'est pas encore clair. Une étude récente des protéines qui interagissent avec Xist a permis l'identification de 200 partenaires qui appartiennent à plusieurs classes fonctionnelles (modification des histones ou de l'ADN, topologie de l'ADN, l'organisation de la chromatine

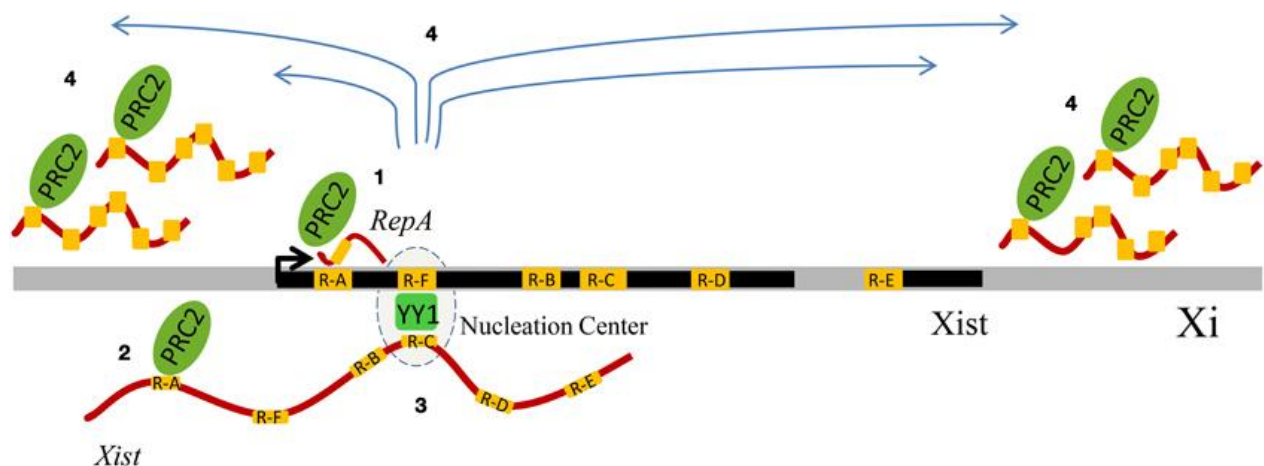




# Introduction

et des chromosomes, l'interaction avec la matrice et l'enveloppe nucléaire, ...) illustrant la complexité de ce mode de dérégulation (Minajigi *et al*, 2015).

L'étude de l'inactivation du chromosome X a permis de comprendre beaucoup de choses sur l'action en *cis* d'un ARN non codant. Mais des découvertes plus récentes ont montré que l'expression de Xist était elle-même régulée par d'autres ARN non codants tels que Tsix ou Jpx. Cela fait de ce mécanisme un modèle important pour notre compréhension des ARN non codants et met en avant la possibilité de réseaux de régulation impliquant des ARN non codants.



**Figure 5 : Schéma du mode d'action de l'ARN non codant Xist** (Casa et Gabellini, 2012).

La séquence RepA sur l'ARN Xist recrute le complexe de répression Polycomb PRC2. Grâce à une autre séquence en son sein, « Repeat C », l'ARN Xist s'associe avec la protéine YY1 ce qui permet la fixation au « nucleation center » et la mise en place en *cis* de la répression.

## L'action en trans de l'ARN HOTAIR

HOTAIR ou « HOX Transcript Antisense Intergenic RNA » est un ARN non codant impliqué dans la régulation des gènes du développement HOX. Ce dernier, décrit pour la première fois en 2007 (Rinn *et al*, 2007 ; Woo et Kingston, 2007), est un ARN de 2,2 kilobases, épissé (6 exons) et polyadénylé. Il est transcrit à partir du chromosome 12 (12q13.13) dans le locus HOXC sur le brin antisens de la région intergénique entre les gènes HOXC 11 et HOXC 12. L'ARN HOTAIR est co-exprimé avec l'ensemble du locus HOXC, son profil d'expression est très proche de celui du gène HOXC 11.



# Introduction

---

Malgré sa présence dans le locus HOXC, le silencing de HOTAIR n'a pas d'effet sur les gènes du locus mais permet la dérégulation du locus HOXD qui est situé sur le chromosome 2 (Rinn *et al*, 2007) (Fig. 6). Depuis sa découverte, l'ARN HOTAIR a été beaucoup étudié du fait que sa surexpression a été observée dans de nombreux cancers et constitue un marqueur du potentiel métastatique. Ainsi HOTAIR fait partie des gènes « driver » du cancer et constitue un des exemples de participation d'un ARN non codant à une pathologie (Gupta *et al*, 2010 ; Geng *et al*, 2011 ; Nakagawa *et al*, 2013).

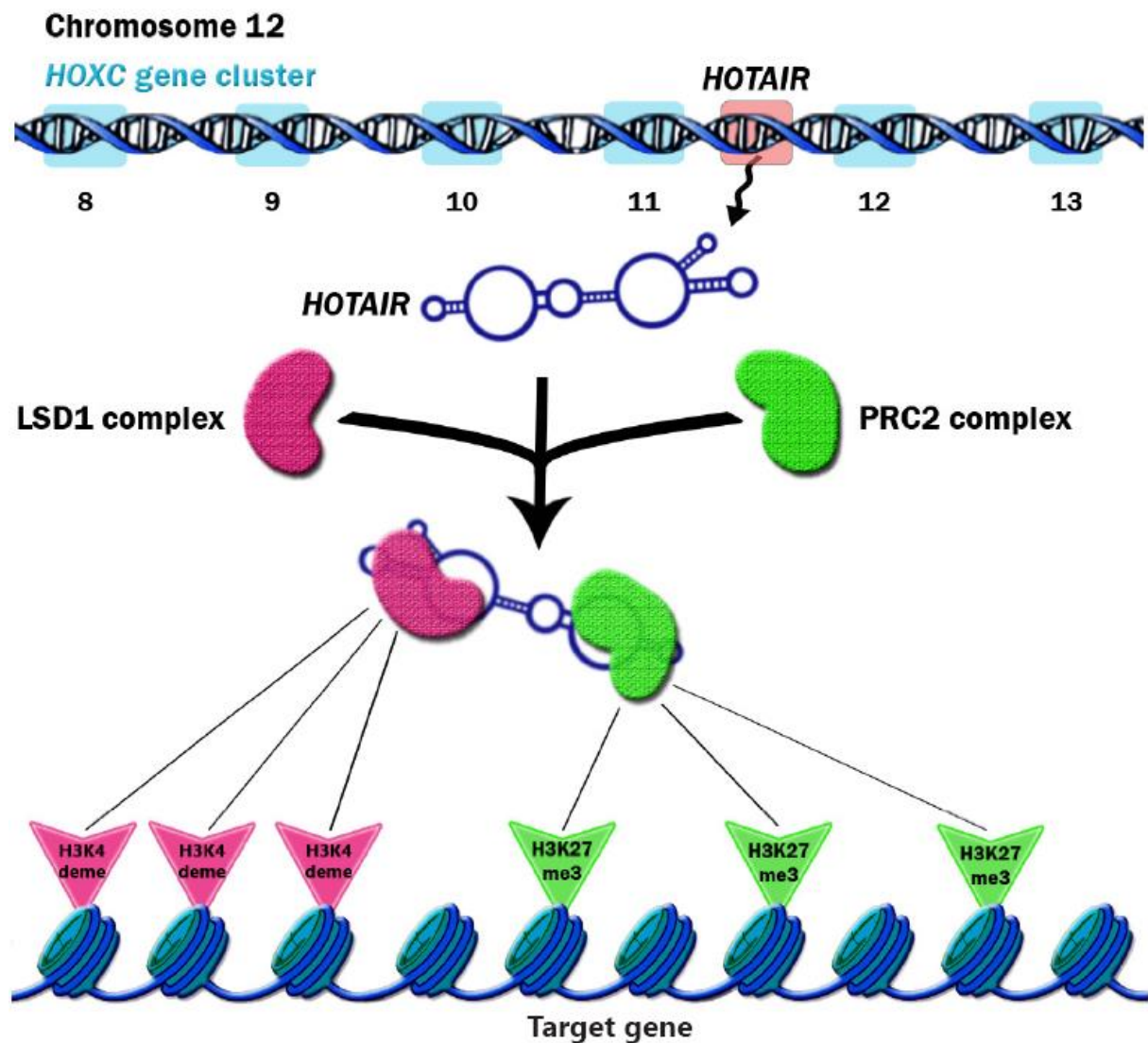
Au niveau du mécanisme d'action, de la même façon que l'ARN Xist, HOTAIR est capable de recruter des complexes modificateurs d'histones. Le recrutement du complexe répresseur Polycomb PRC2 qui apporte la méthylase d'histone EZH2 induit la triméthylation de H3K27 (Tsai *et al*, 2010). Le complexe PRC2 interagit avec HOTAIR grâce à une structure en tige-boucle située en 5'. Mais l'ARN HOTAIR possède également une seconde structure cette fois-ci en 3' qui permet le recrutement d'un second complexe protéique, LSD1 (Tsai *et al*, 2010). LSD1 est une histone déméthylase qui va permettre de retirer la marque de l'euchromatine H3K4me3. Ceci permet à HOTAIR d'apporter sur le locus HOXD les complexes permettant de retirer les marques d'activation et d'ajouter les marques du silencing sur les histones (Fig. 6). Le recrutement de plusieurs complexes sur un ARN non codant est fréquemment observé pour les ARN non codants exprimés durant le développement. On retrouve ainsi plus de 30 ARN non codants exprimés dans les cellules souches embryonnaires qui interagissent avec plusieurs complexes modificateurs chromatinien permettant l'effacement ou l'écriture des marques épigénétiques (Guttman *et al*, 2011). Les ARN non codants pourraient ainsi coordonner l'activité des complexes protéiques en introduisant un niveau d'organisation au-delà de celui de la formation des complexes.

Le mode d'interaction entre HOTAIR et le gène cible du locus HOXD n'est pas complètement identifié. Sur la base des mécanismes observés pour d'autres ARN on peut proposer que : (i) Comme pour Xist, HOTAIR interagit avec une protéine de liaison à l'ADN qui va lui permettre de se fixer à un site spécifique dans le locus ciblé. (ii) Il peut également interagir directement avec l'ADN par une hybridation ADN/ARN ou par formation d'un triplex ADN double brin – ARN. C'est le cas en particulier de certains ARN non codants interagissant spécifiquement avec les promoteurs des ribosomes (Schmitz *et al*, 2010). (iii) Les transcrits naissants peuvent recruter les complexes régulateurs et ainsi agir sur les gènes voisins. Ce dernier cas de figure



# Introduction

devrait *a priori* s'appliquer uniquement aux actions en *cis*. Cependant, il est possible que les longs ARN non codants profitent de l'organisation 3D de la chromatine dans le noyau pour agir sur leurs cibles. Ainsi dans le cas des gènes HOX, il a été montré que malgré leur présence sur différents chromosomes, ils étaient souvent très proches au niveau spatial dans le noyau (Denholtz *et al*, 2013 ; Noordermeer *et al*, 2014).



**Figure 6 : Schéma du mode d'action de l'ARN non codant HOTAIR** (Hajjari et Salavaty, 2015).

L'ARN HOTAIR est transcrit à partir du locus HOXC sur le chromosome 12 mais agit en *trans* sur le locus HOXD situé sur le chromosome 2. L'ARN HOTAIR possède une structure secondaire qui permet le recrutement de deux complexes modificateurs d'histone. Le complexe LSD1 qui permet de retirer les marques activatrices de la transcription H3K4me3 et le complexe PRC2 qui permet la pose de la marque répressive de la transcription H3K27me3.

## 2.2.2 – Interaction avec les facteurs de transcription

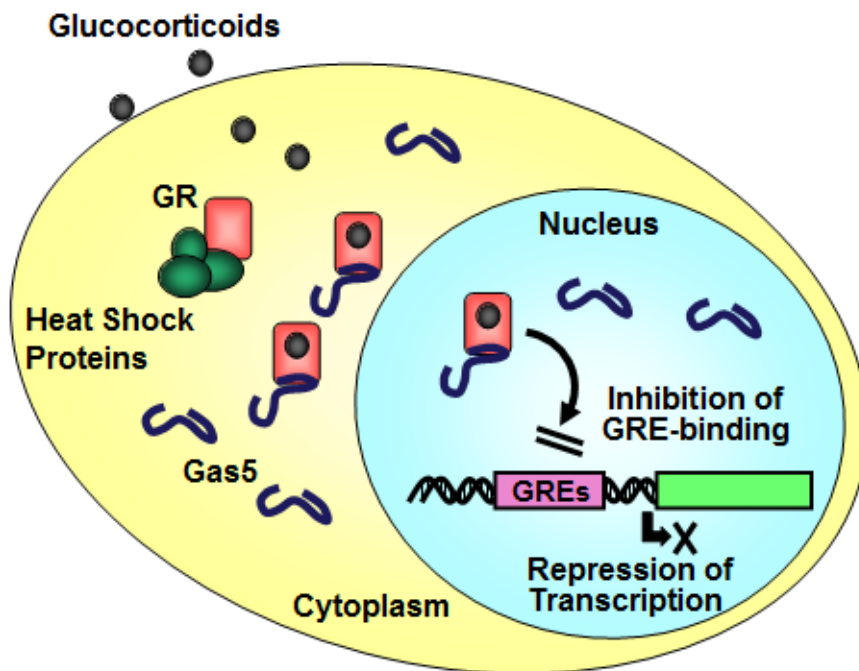
Les modifications de la chromatine vont avoir un impact sur la transcription en modifiant l'accessibilité aux facteurs de transcription. Cependant l'ARN non codant lui-même peut interagir directement avec ceux-ci comme il a été observé pour l'ARN non codant GAS5 ou « Growth-arrest specific 5 ». Le gène GAS5 contient 12 exons mais ne code aucune protéine (Smith *et al*, 1998). 10 petits ARN nucléolaires sont produits à partir des introns de GAS5 tandis que les deux isoformes épissées génèrent les longs ARN non codants GAS5a et GAS5b ; GAS5a étant l'isoforme majoritaire. Cet ARN non codant est impliqué dans la régulation des gènes de réponse aux glucocorticoïdes. Ces gènes sont positivement régulés par les récepteurs des glucocorticoïdes (GRs) qui appartiennent à la superfamille des facteurs de transcription hormone-dépendant. Ces facteurs de transcription sont cytoplasmiques, mais lorsqu'ils se lient à un agoniste des glucocorticoïdes ils migrent vers le noyau et interagissent avec les « glucocorticoid response elements » (GREs) sur l'ADN (Chrousos *et al*, 2005). Les gènes de réponse aux glucocorticoïdes sont impliqués dans de nombreuses fonctions cellulaires dont la prolifération et des mécanismes anti-apoptotiques. Aussi, comme GAS5 inhibe leur expression, est-il considéré comme un gène suppresseur de tumeur (à l'opposé de HOTAIR qui appartient à la catégorie des oncogènes). Comme l'indique son nom (Growth-arrest specific 5), GAS5 a été initialement décrit du fait de son expression dans des cellules rendues quiescentes par privation soit de nutriments soit de facteurs de croissance (Schneider *et al*, 1988). Le mode d'action du long ARN non codant GAS5 passe par son interaction avec le récepteur des glucocorticoïdes (Fig. 7). L'ARN GAS5 sera capable via des structures secondaires en tige-boucle à son extrémité 3' de mimer un « glucocorticoid response element ». Cette séquence appelée « GRE-mimic » va alors leurrer les récepteurs des glucocorticoïdes et les détourner de leurs gènes cibles (Kino *et al*, 2010). Dans la même étude, il est montré qu'il existe dans cette même région en 3' de l'ARN GAS5 une autre structure secondaire permettant de mimer un « mineralocorticoide response element » signe que cet ARN pourrait également interagir avec d'autres facteurs de transcription hormone-dépendant.

Il est aussi intéressant de noter que des piARN peuvent être générés à partir des snoRNA contenus dans les introns de GAS5. Ces piARN induisent l'expression de plusieurs gènes dont TRAIL qui peut induire l'apoptose dans les cellules tumorales. L'activité de gène suppresseur



# Introduction

de tumeur de GAS5 peut donc résulter de l'activité de plusieurs ARN non codants issus de ce gène. Dans le cas de ces piARN, la stimulation de l'expression se fait via le recrutement du complexe MLL3/COMPASS qui apporte l'activité méthyltransférase de H3K4 et déméthylase de H3K27. Ainsi, le même gène utilise aussi bien des mécanismes au niveau de la chromatine que des facteurs de transcription.



**Figure 7 : Schéma du mode d'action de l'ARN non codant GAS5** (Monteiro *et al*, 2013).

L'ARN GAS5 est capable de mimer un élément de réponse aux glucocorticoïdes via une structure secondaire en son extrémité 3'. Cela va leurrer les récepteurs aux glucocorticoïdes qui vont se fixer sur l'ARN GAS5 et donc ne pourront pas activer la transcription de leurs gènes cibles.

D'autres longs ARN non codants ont été décrits comme pouvant interagir avec des facteurs de transcription et cette interaction ne mène pas toujours à une répression des gènes cibles de ces facteurs de transcription. C'est par exemple le cas de l'ARN non codant Evf2 qui par son expression peut permettre à la fois la répression mais aussi l'induction de l'expression de gènes avoisinants. Evf2 est un ARN de 3,8 kilobases, épissé et polyadénylé, transcrit sur le brin antisens d'un enhancer très conservé situé dans la région intergénique entre les gènes du développement neuronal *Dlx-5* et *Dlx-6* (Feng *et al*, 2006) (Fig. 8). L'augmentation de l'expression de cet ARN est associée à la différenciation des progéniteurs des neurones



# Introduction

---

GABAergiques et a également été observée chez les cellules souches embryonnaires (ES) chez les souris (Dinger *et al*, 2008 ; Mercer *et al*, 2010). L'expression de l'ARN Evf2 est sous la dépendance du facteur de croissance Sonic Hedgehog (SHH) (Feng *et al*, 2006).

La fonction de l'ARN Evf2 est de moduler l'activité de l'enhancer de la région intergénique dont il est issu et qui contrôle l'expression de *Dlx-5* et *Dlx-6*. En s'associant avec des facteurs de transcription, Evf2 facilite leur fixation aux séquences de l'enhancer. Les facteurs interagissant avec l'ARN non codant varient selon le stade du développement ou de différenciation et activer ou réprimer l'activité enhancer. Un dysfonctionnement de ces gènes régulateurs du devenir des neurones GABAergiques a été observé dans de nombreuses maladies psychiatriques comme la schizophrénie, l'autisme ou le syndrome de Gilles de la Tourette (Di Cristo *et al*, 2007). Pour augmenter l'activité transcriptionnelle, l'ARN Evf2 forme un complexe avec la protéine à homéodomaine Dlx-2. Dlx-2 interagit avec l'ARN via une séquence conservée située en 5' (Feng *et al*, 2006) et le complexe formé pourra se fixer à des séquences spécifiques de la région enhancer via l'homéodomaine de Dlx-2 et stimuler l'activité transcriptionnelle des gènes *Dlx-5* et *Dlx-6* (Fig. 8). L'ARN Evf2 est essentiel pour que la protéine Dlx-2 se fixe à la région enhancer et aucune fixation de Dlx-2 n'est observée en l'absence de l'ARN Evf2 (Bond *et al*, 2009). Cependant, un arrêt prématuré de la transcription de Evf2 induit par l'introduction de séquences de polyadénylation et qui devrait conduire à une perte de fonction (et donc de l'activité de l'enhancer) induit une augmentation de l'expression des gènes *Dlx-5* et *Dlx-6*. Ceci a permis de mettre en évidence un second partenaire de l'ARN Evf2, MECP2, une protéine de liaison à l'ADN méthylé associée à la répression de la chromatine par sa capacité à recruter des histones désacétylases (HDAC) (Nan *et al*, 1997). L'ARN Evf2 va ainsi pouvoir réguler positivement ou négativement l'expression de ces gènes (Fig. 8).

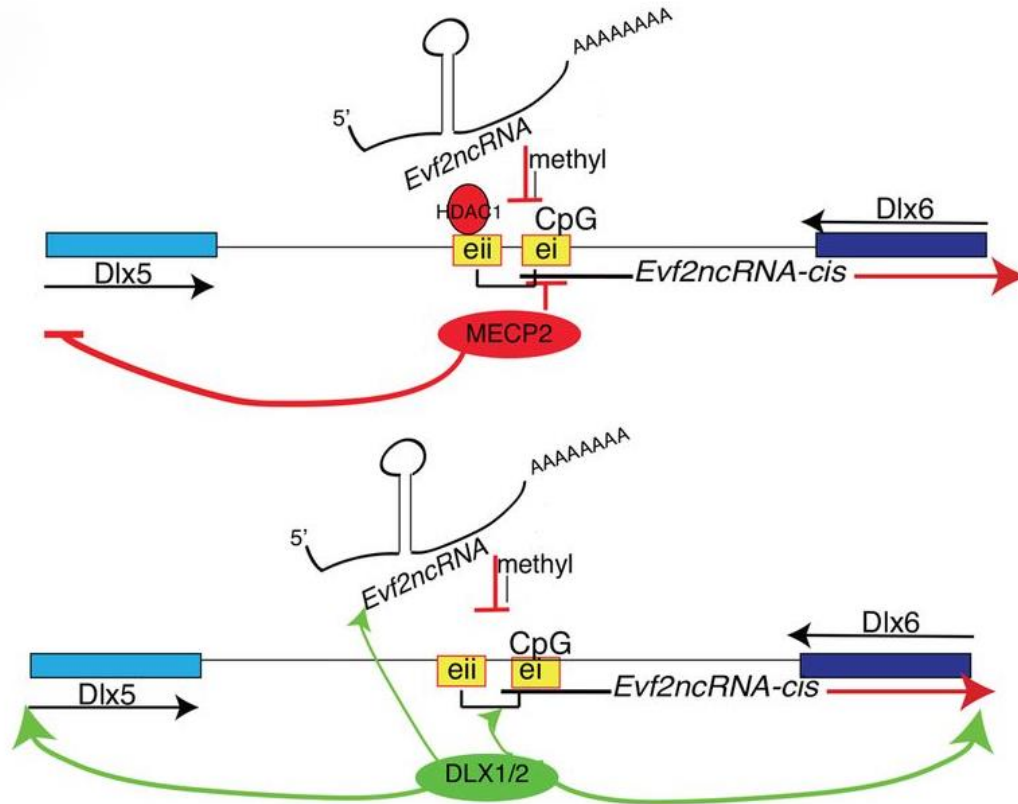
Malgré un mode d'action qui semble être en *cis*, l'action peut également se faire en *trans*. En effet en absence de l'ARN Evf2, l'introduction dans les lignées d'une construction plasmidique permettant l'expression de l'ARN stimule à nouveau l'induction de l'expression des gènes du locus (Bond *et al*, 2009). Il est également observé dans ces lignées que l'expression du gène *Dlx-6* est plus fortement induite que celle du gène *Dlx-5*. Cela suggère que l'ARN Evf2 qui recouvre en partie le gène *Dlx-6* pourrait avoir un effet inhibiteur direct sur l'expression du gène sur le brin antisens à travers un conflit transcriptionnel en *cis* (voir





# Introduction

plus loin) (Feng *et al*, 2006). Ainsi l'ARN Evf2 peut agir de plusieurs façons à la fois, comme un co-activateur ou co-répresseur de la transcription en *trans*, mais également en *cis* sur *Dlx-6* comme un ARN antisens.



**Figure 8 : Schéma du mode d'action de l'ARN non codant Evf2** (Berghoff *et al*, 2013).

L'ARN *Evf2* est issu de la région enhancer située entre les gènes *Dlx-5* et *Dlx-6* et qui contrôle leur expression. L'ARN *Evf2* est capable de moduler l'activité de la région enhancer en fonction du stade de développement en s'associant avec deux facteurs de transcription. Ainsi selon le stade de développement l'ARN *Evf2* soit s'associe avec le facteur de transcription *Dlx-2* pour activer l'expression des gènes *Dlx-5* et *6*, soit s'associer avec l'élément répresseur *MECP2* qui désacétyle les histones et empêcher l'expression du locus.

## 2.2.3 – Régulation post-transcriptionnelle

Même si la plupart des longs ARN non codants étudiés à ce jour agissent de manière directe ou indirecte sur la transcription, il existe aussi des données sur des actions au niveau post-transcriptionnel. Les possibilités d'action sont variées et concernent la stabilité de l'ARN messager, l'inhibition de la traduction ou une modulation de l'épissage. Ces effets ne sont par ailleurs pas tous négatifs car des ARN non codants peuvent permettre la stabilisation de certains messagers.



## **Dégradation de l'ARN messenger**

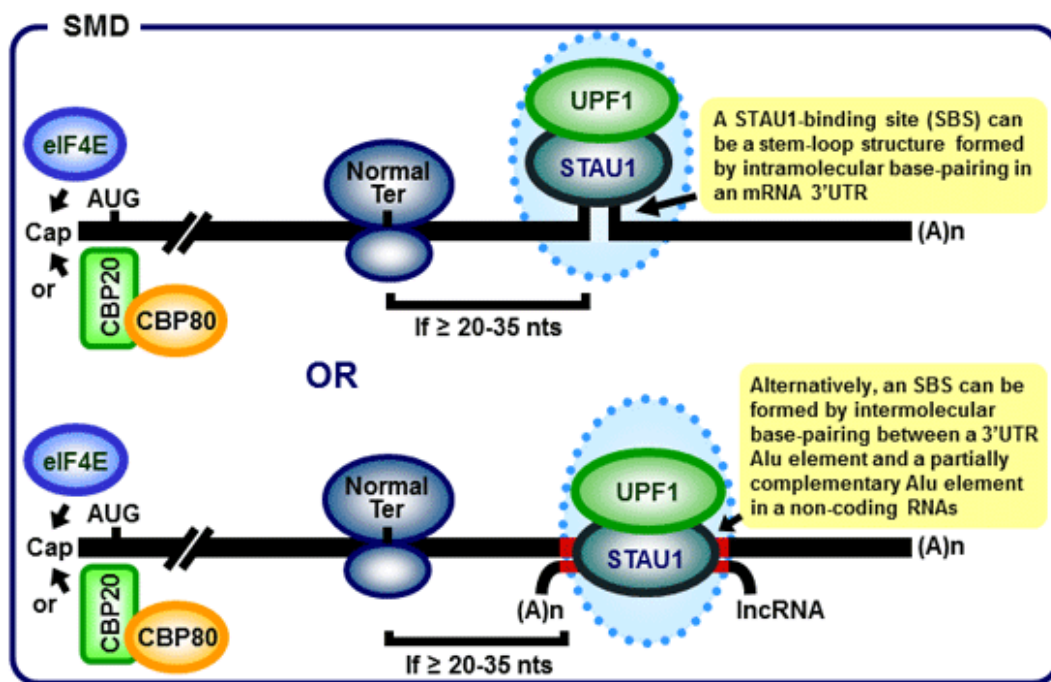
Le devenir des molécules d'ARN messenger est contrôlé à de multiples niveaux. Il existe en particulier plusieurs mécanismes de contrôle qualité comme les mécanismes NMD (Nonsense-mediated mRNA decay), NSD (non-stop mRNA decay) et NGD (no-go mRNA decay) qui induisent la dégradation des transcrits aberrants (du point de vue de l'organisation de la région codante). Néanmoins, ces mécanismes semblent également impliqués dans la régulation de certains ARN messagers « normaux ». L'absence du mécanisme NMD induit une augmentation de l'expression de 10% des gènes chez les mammifères (Mendell *et al*, 2004). Découvert plus récemment, un autre mécanisme permet également la régulation des ARN en cours de traduction, le SMD pour « Staufen1-mediated mRNA decay » (Kim *et al*, 2005). Ce mécanisme, même s'il n'est pas impliqué dans le contrôle qualité mais dans la régulation de certains ARN messagers, agit avec un des partenaires de la voie NMD, la protéine UPF1 (Kim *et al*, 2005). C'est la protéine STAU1 (Staufen1) qui permet le recrutement de la protéine UPF1, jouant ainsi un rôle analogue à celui de l'EJC (Exon junction complex) dans le mécanisme NMD où il recrute UPF1 sur l'ARN messenger en cours de traduction (Fig. 9). STAU1 est une « RNA binding protein » qui interagit avec l'ARN double brin via 4 domaines de liaison à l'ARN double brin (Wickham *et al*, 1999). Des sites de fixation « STAU1 binding site » ont été décrits dans les régions non traduites 3' des ARN ciblés par ce mécanisme. Le site le mieux caractérisé est celui de l'ARN messenger du gène *Arf1* qui comprend une tige de 19 paires de bases et une boucle de 100 nucléotides. La protéine STAU1 va se fixer spécifiquement sur cette structure du 3'UTR et permettre le recrutement d'UPF1 qui va servir de plateforme pour recruter les complexes de décoiffage de l'ARN messenger (Lykke-Andersen *et al*, 2002) (Fig. 9). Même si cette séquence de 19 paires de bases semble être conservée chez d'autres espèces comme le rat et la souris, celle-ci n'a pas été retrouvée chez d'autres cibles connues du mécanisme SMD (Kim *et al*, 2007). Par contre, la fixation de STAU1 aux autres gènes cibles nécessite l'action d'un ARN non codant appelé 1/2 sbs RNA1. Cet ARN polyadénylé de 388 nucléotides a la particularité de contenir dans sa séquence celle de l'élément transposable Alu. L'élément Alu est présent à environ 1,4 million de copies dans le génome humain (Batzer *et al*, 2002) et des copies sont ainsi retrouvées dans les régions 3'UTR de nombreux ARN messagers ciblés par le





# Introduction

mécanisme SMD. C'est le cas par exemple de l'ARN messager de la protéine SERPINE1, une des cibles de  $\frac{1}{2}$  sbs RNA1 (Gong et Maquat, 2011). Ainsi les éléments Alu en 3'UTR de l'ARN messager et la séquence de l'ARN  $\frac{1}{2}$  sbs RNA vont pouvoir s'hybrider et reconstituer la séquence permettant le recrutement de STAU1 (Fig. 9). L'action d'un siARN contre  $\frac{1}{2}$  sbs RNA1 montre une augmentation de SERPINE1 de la même valeur qu'un siARN ciblant STAU1, signe que l'ARN non codant est indispensable à la fixation de STAU1 (Gong et Maquat, 2011). Il a été calculé qu'environ 350 longs ARN non codants possédaient au moins un élément Alu dans leur séquence. Il a aussi été observé que  $\frac{1}{2}$  sbs RNA pouvait agir sur plusieurs cibles et qu'un ARN messager pouvait être ciblé par plusieurs  $\frac{1}{2}$  sbs RNA. Cela fait du SMD un mécanisme de régulation post-transcriptionnel puissant. Néanmoins, tous les ARN messagers possédant un élément Alu dans leurs séquences ne vont pas permettre une régulation via ce mécanisme car la fixation de la protéine STAU1 peut être empêchée par d'autres ribonucléoprotéines (Gong *et al*, 2012).



**Figure 9 : Schéma du mécanisme SMD, Staufen1-mediated mRNA Decay (Maquat *et al*, 2011).**

Deux mécanismes permettent la mise en place du SMD. Le premier est que la protéine STAU1 (Staufen1) reconnaisse une tige-boucle spécifique de 19 paires de base située dans la région non traduite 3' de l'ARNm. Ceci est le cas spécifique de l'ARNm du gène Arf1. Le second est que l'ARN non codant  $\frac{1}{2}$  sbs RNA reconnaisse une séquence Alu dans la région non traduite 3' et crée ainsi un site de fixation pour STAU1. Dans les deux cas, la fixation de STAU1 permet la fixation de la protéine UPF1 qui permet le recrutement de facteurs qui décoifferont l'ARNm cible qui sera alors dégradé par les exonucléases de la cellule.



## ***Inhibition de la traduction***

L'inhibition de la traduction est un mécanisme bien connu, en particulier, dans le cadre de la régulation par les microARN. Mais ce mode d'action n'est pas réservé aux petits ARN interférants et il existe au moins un cas d'un long ARN permettant l'inhibition de la traduction de ses cibles, le long ARN intergénique p21. Cet ARN de 3,1 kilobases, épissé (2 exons), est transcrit sur le brin antisens dans une région intergénique située à 15 kilobases en amont du gène p21. Son expression est dépendante du facteur de transcription p53 et sa fonction première est de réprimer de nombreux gènes régulés par l'expression de p53 (Huarte *et al*, 2010). Ces cibles sont pour beaucoup impliquées dans la réponse anti-apoptotique et l'expression du long ARN intergénique p21 va promouvoir l'apoptose contribuant ainsi à l'activité biologique de p53. Le mode de reconnaissance de ses cibles est encore mal connu mais pourrait utiliser une hybridation en plusieurs sites avec l'ARN messager. Ainsi on peut observer dans la séquence du transcrit  $\beta$ -caténine, une des cibles du long ARN intergénique p21, 15 sites partiellement complémentaires de la séquence du lincRNA-p21 (motifs d'une vingtaine de nucléotides comprenant au moins 13 nucléotides appariés) (Yoon *et al*, 2012).

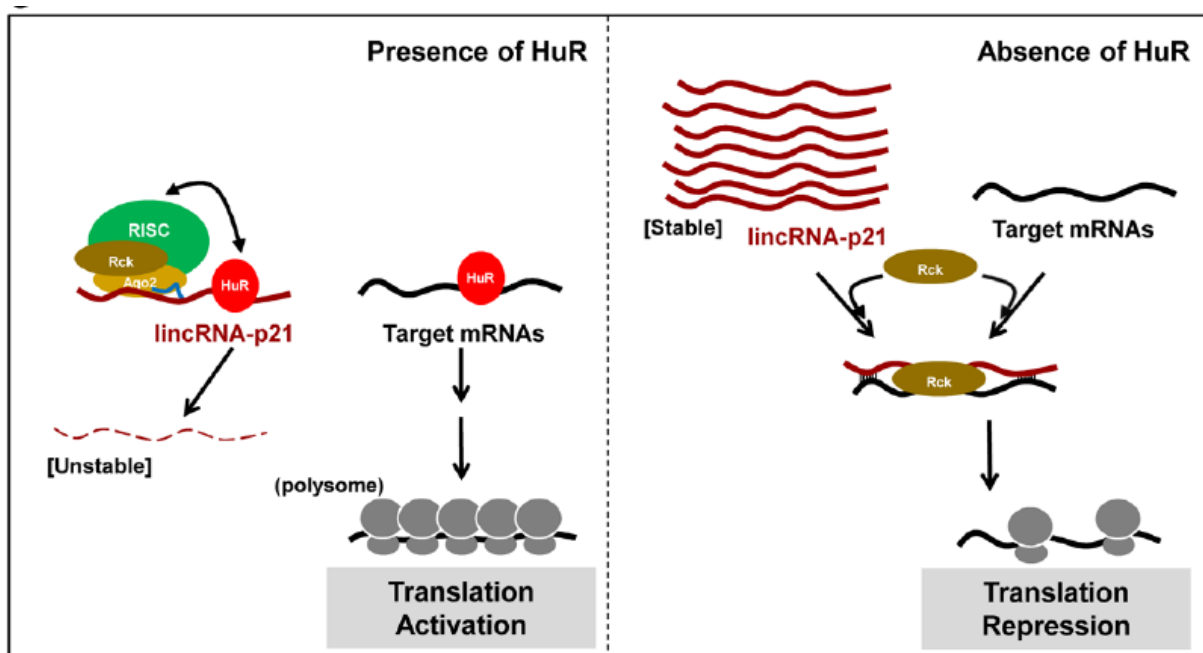
La même étude a mis en évidence que la régulation par cet ARN non codant passait par le recrutement de Rck/p54. Rck/p54 est connu pour être un inhibiteur de la traduction (Coller *et al*, 2005). Il a d'ailleurs été montré que Rck/p54 était impliqué dans l'inhibition de la traduction dans le mécanisme miARN. Cette hélicase à boîte DEAD interagit avec les protéines Ago1 et Ago2 dans la voie miARN et facilite la formation des P-bodies. Le silencing de Rck/p54 entraîne une perte de la capacité d'inhibition de la traduction du complexe miRISC (Chu et Rana, 2006).

Au vu de son effet pro-apoptotique, l'expression de l'ARN intergénique p21 est très régulée. Cette régulation va se faire au niveau post-transcriptionnel par la « RNA binding protein HuR » (Yoon *et al*, 2012). HuR est une protéine ubiquitaire qui influe sur beaucoup de mécanismes cellulaires en agissant sur la stabilité et en modulant la traduction (Hinman et Lou, 2008 ; Abdelmohsen et Gorospe, 2010). La protéine HuR agit sur le devenir des messagers ciblés en coopérant ou en entrant en compétition avec les protéines de la voie de dégradation des ARN (Young *et al*, 2009). Dans le cas de l'ARN intergénique p21, la protéine



# Introduction

HuR va déstabiliser le transcrit et conduire à sa dégradation via une association avec la protéine Argonaute et le complexe RISC (Yoon *et al*, 2012). Ce phénomène avait déjà été observé dans la dégradation du messenger c-Myc par HuR qui est facilitée par l'interaction avec le complexe RISC (Kim *et al*, 2009). Ainsi l'activité de l'ARN non codant p21 sera dépendante de l'action de la protéine HuR. En sa présence l'ARN sera dégradé et ne pourra donc pas avoir d'action sur les transcrits cibles (Fig. 10). En absence de la protéine HuR, l'ARN intergénique p21 est stable et peut interagir avec ses cibles. Le recrutement sur la cible de Rck/p54 permettra alors la répression de la traduction (Yoon *et al*, 2012) (Fig. 10). La régulation des ARN non codants par HuR associé à let7/Ago2 ne semble pas limitée à ce cas et, par exemple, intervient aussi dans la dégradation du long ARN non codant HOTAIR (Yoon *et al*, 2014).



**Figure 10 : Schéma du mode d'action du long ARN intergénique p21** (Yoon *et al*, 2012).

Le long ARN intergénique p21 est dégradé par la présence de la protéine régulatrice HuR. HuR se fixe sur l'ARN intergénique p21 et en association avec le complexe RISC permet la dégradation de l'ARN. En son absence, le long ARN intergénique p21 va pouvoir s'associer avec Rck/p54 qui permet l'inhibition de la traduction des ARNm cibles.

## Modulation de l'épissage

L'épissage alternatif est une étape clef à la fois dans la régulation et la diversité de fonction de nombreux gènes. Il est estimé que plus de 95% des gènes contenant des introns sont

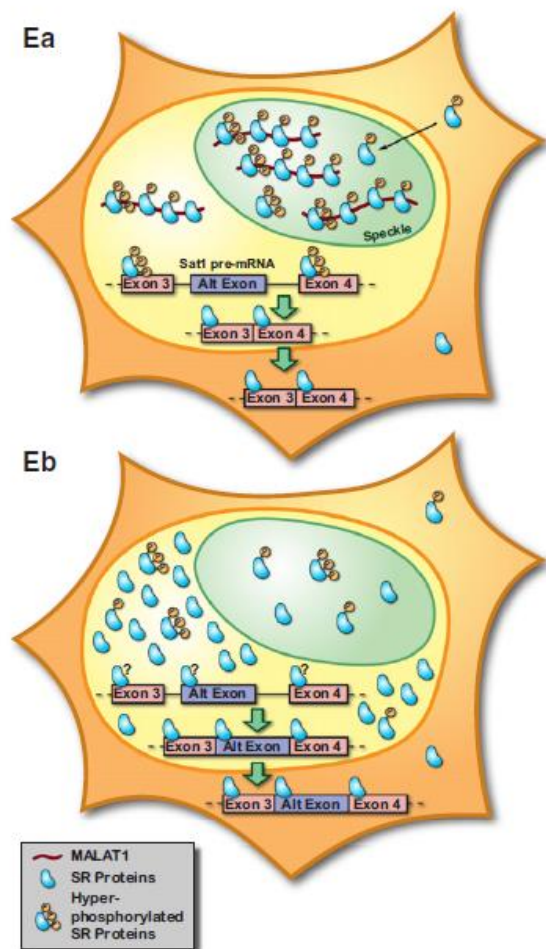


soumis à des épissages alternatifs et que les variants d'épissage produits diffèrent selon le type cellulaire (Pan *et al*, 2008 ; Wang *et al*, 2008). Le complexe d'épissage ou « Spliceosome » est composé de nombreux partenaires. Les premiers sont les petites ribonucléoprotéines nucléaires ou snRNP (small nuclear ribonucleoprotein). Ce sont des complexes formés par les petits ARN nucléaires et leurs protéines associées et qui participent directement à l'épissage. Il existe aussi plus d'une centaine de protéines qui s'associent avec ces complexes dont, par exemple, les protéines de la famille SR (protéines qui doivent leur nom à un domaine riche en sérine (S) et arginine (R)). Les protéines SR sont des protéines de liaison à l'ARN dont le rôle est essentiel dans l'épissage conventionnel comme alternatif (Blencow *et al*, 2000). Dans l'épissage alternatif, ces protéines sont importantes pour la reconnaissance des séquences activatrices d'épissage présentes dans les exons (exonic splicing enhancer, ESE) situées dans des exons alternatifs (Lin et Fu, 2007 ; Long et Caceres, 2009). Le niveau en protéine SR est finement régulé et la modification de concentration ou de phosphorylation de ces protéines peut ainsi modifier l'épissage de nombreux pré-messagers (Bourgeois *et al*, 2004). Cette régulation dépend entre autres du long ARN non codant MALAT1 (Metastasis-Associated Lung Adenocarcinoma Transcript 1). Ce transcrit de 7 kilobases est dépourvu d'intron et non polyadénylé. Historiquement identifié à cause de son absence d'expression dans plusieurs cancers (Ji *et al*, 2003 ; Lin *et al*, 2006), sa fonction a longtemps été inconnue. C'est sa localisation abondante dans les « nuclear speckles », sites d'accumulation de la machinerie d'épissage dans le noyau, qui a suggéré sa fonction. MALAT1 interagit avec de nombreux facteurs d'épissage mais ne joue aucun rôle dans la formation ou la maintenance de la structure des « nuclear speckles », au contraire de l'ARN non codant NEAT1. Le rôle de MALAT1 est de réguler la distribution des facteurs d'épissage dans cette structure (Tripathi *et al*, 2010). Il a été montré par la même équipe que la délétion de MALAT1 a pour conséquence une délocalisation des facteurs d'épissage (Fig. 11), qui vont être retrouvés en plus grande quantité en dehors des « nuclear speckles ». Cela va également amener des changements dans l'épissage alternatif. De plus, l'ARN MALAT1 est important pour la phosphorylation des protéines SR (Tripathi *et al*, 2010). L'état de phosphorylation des protéines SR influe sur leur activité et un cycle de phosphorylation/déphosphorylation est nécessaire pour l'épissage. Ainsi une hyperphosphorylation des SR protéines est nécessaire pour la fixation à l'ARN pré-messager



# Introduction

et donc le recrutement du spliceosome. Au contraire, une déphosphorylation partielle sera nécessaire pour la transestérification, étape clef de l'épissage (Cao *et al*, 1997). Il est possible que l'ARN MALAT1 module la phosphorylation des protéines SR en interagissant avec les kinases (SRPKs) ou les phosphatases (PP1 ou PP2A) qui les modifient. Des données ont montré que SRPK1 et PP1 pouvaient influencer sur les épissages alternatifs (Shi et Manley, 2007 ; Zhong *et al*, 2009) et la kinase SRPK1 est délocalisée hors des « nuclear speckles » en absence de l'ARN MALAT1 (Tripathi *et al*, 2010). Le spectre d'action de MALAT1 en agissant sur les protéines SR ne semble pas se limiter aux mécanismes d'épissage mais s'étend également au transport des ARN. La délétion de MALAT1 va ainsi augmenter la concentration de protéines SRSF1 déphosphorylées, marques des ARN messagers à exporter. On retrouve ainsi une proportion d'ARN polyadénylé plus grande dans le cytoplasme (Tripathi *et al*, 2010).



**Figure 11: Mécanisme de régulation par l'ARN non codant MALAT1** (Tripathi et al, 2010). L'ARN MALAT1 est impliqué dans la régulation de l'épissage. La présence de l'ARN MALAT1 permet une colocalisation de l'ensemble des facteurs, telles que les protéines SR, les phosphatases et kinases de ces protéines SR, permettant l'épissage au sein des « nuclear speckles » siège du mécanisme d'épissage. En son absence, les protéines SR ne sont plus phosphorylées ne permettant plus la fixation du spliceosome.



## ***Stabilisation du transcrit***

Tout comme pour les autres activités des longs ARN non codants, les effets au niveau post-transcriptionnel ne sont pas toujours inhibiteurs de l'expression. Plusieurs cas de stabilisation de l'ARN messenger cible ont ainsi été bien décrits. Le cas le plus répandu est celui des « éponges à miARN » aussi appelées ceRNA (Competing endogenous RNA). Le modèle par excellence de cette classe d'ARN non codant est le pseudogène PTENP1 (Tay *et al*, 2011). PTENP1 est un pseudogène du gène suppresseur de tumeur PTEN, et bien que le cadre de lecture de la protéine PTEN n'ait pas été maintenu, la séquence est globalement conservée en particulier dans la région non traduite 3'. Du fait de cette conservation en 3', les deux ARN possèdent les mêmes sites pour les miARN et sont régulés par les mêmes microARN (Tay *et al*, 2011). L'expression de PTENP1 a donc pour conséquence de mobiliser les miARN ciblant le messenger PTEN et ainsi de détourner la régulation par ces microARN de l'ARNm PTEN vers l'ARN PTENP1. L'expression de l'ARN PTENP1 permet ainsi une meilleure expression de PTEN tandis que la perte du pseudogène diminue l'expression de PTEN ce qui aura pour conséquence d'activer la voie PI3K/AKT et permettre la prolifération (Tay *et al*, 2011). Au même titre que PTEN, PTENP1 est donc un suppresseur de tumeur et la diminution ou la perte de son expression est observée dans de nombreux cancers. D'autres pseudogènes sont probablement impliqués dans la régulation du gène parental par ce type de compétition, c'est par exemple le cas du pseudogène KRAS (Poliseno *et al*, 2010).

D'autres types de mécanismes permettant la stabilisation du transcrit cible existent. C'est le cas de l'ARN BACE1-AS. Cet ARN de 2 kilobases est transcrit à partir du brin antisens de la région intronique du gène BACE1. Ce dernier recouvre ainsi la totalité de l'exon 6 de BACE1 (Faghihi *et al*, 2008). Ce long ARN non codant est impliqué en particulier dans la maladie d'Alzheimer ou son expression est élevée. En agissant sur l'expression de BACE1, l'ARN BACE1-AS va favoriser dans le cerveau la formation des agrégats de peptides amyloïdes responsables de la maladie. En effet, BACE1 code la  $\beta$ -secretase1 qui permet la première coupure des précurseurs amyloïde (AAP) (Vassar *et al*, 1999). L'ARN BACE1-AS va stabiliser l'ARN messenger BACE1 en empêchant la fixation du miARN miR-485-Sp du fait de l'occlusion du site de fixation par la formation d'un duplex ARN messenger – ARN non codant (Faghihi *et al*, 2010). Cette formation d'un duplex ARN-ARN avec un ARN messenger se retrouve chez



d'autres ARN non codants et conduit à la stabilisation de l'ARN messager ciblé. L'ARN Wrap53 permet ainsi de stabiliser l'ARN messager p53. L'exon 1 de l'ARN Wrap53 recouvre l'exon 1 de l'ARN p53 ce qui permet aux deux transcrits de s'hybrider sur 227 paires de bases. Le duplex Wrap53/p53 en tête à tête va empêcher la dégradation de l'ARN p53 et favorisera ainsi son expression en particulier dans le cas de réponse au dommage à l'ARN (Mahmoudi *et al*, 2009).

## 3 – Les longs ARN antisens

### 3.1 – Généralités et spécificités

Les ARN antisens constituent une classe particulière de longs ARN non codants. Découverts initialement chez les bactéries il y a plus de 30 ans, la question de leur importance biologique chez les eucaryotes est plus récente (Vanhée-Brussollet et Vaquero, 1998). Par définition, les ARN antisens sont les produits de la transcription du brin opposé de gènes codant des protéines. En fonction de l'organisation relative des transcrits sens et antisens il existe probablement différentes classes de transcrits antisens avec des conséquences biologiques spécifiques. L'existence de transcrits antisens est un phénomène fréquent chez les eucaryotes et il a été observé que chez l'homme plus de 63% des ARNm avaient un partenaire antisens (Katayama *et al*, 2005) tandis que 30% des transcrits annotés à ce jour étaient des transcrits antisens (Ozsolak *et al*, 2010).

Ces ARN antisens ou NAT (« Natural Antisense Transcripts) peuvent avoir différentes origines. Ils peuvent venir, comme c'est le cas chez les levures, de l'activité bidirectionnelle des promoteurs qui vont permettre le recouvrement sur le brin complémentaire du gène voisin, lorsque les gènes adjacents sont dans la même orientation (Neil *et al*, 2009). Ils peuvent également avoir pour origine des promoteurs cryptiques présents dans la séquence d'un gène ou profiter de séquences promotrices apportées par un élément transposable (Conley *et al*, 2012).

De la même façon que les longs ARN non codants en général, les NAT vont pouvoir agir en *cis* ou en *trans*. L'origine des *trans*-NAT ne répond pas forcément aux mêmes critères que les *cis*-NAT. Une des hypothèses de leur origine est qu'ils soient issus de pseudogènes (Muro *et al*, 2010). On retrouve le cas par exemple avec le gène PTEN et le pseudogène PTENP1. La transcription antisens des pseudogènes semble être une source majoritaire des *trans*-NATs.





On peut noter que dans le cas de PTEN, en plus d'être régulé positivement par le long ARN non codant PTENP1 issu de son pseudogène, il est régulé mais cette fois-ci négativement par le PTENP1-asRNAα également issu du pseudogène PTENP1 (Johnsson *et al*, 2013).

En plus de cette classification en *cis*-NAT et *trans*-NAT, les transcrits antisens sont classés vis-à-vis de leurs dispositions par rapport au gène sens. On retrouve donc 3 catégories, le « head-to-head » où les transcrits sens et antisens vont se recouvrir sur leurs extrémités 5', le « tail-to-tail » où les transcrits sens et antisens vont se recouvrir sur leurs extrémités 3', et le « l'internal » ou « full overlapping » où le transcrit antisens va recouvrir l'intégralité du gène sens.

Les NAT caractérisés jusqu'à présent sont transcrits par l'ARN Polymérase II et, comme les ARN messagers et les autres longs ARN non codants, ils peuvent être épissés et polyadénylés. Leur niveau d'expression est en général plus faible que celui du sens, avec des transcrits 10 fois moins abondants en moyenne (Ozsolak *et al*, 2010). L'expression de l'antisens est souvent influencée par l'activité du gène sens à travers des effets « enhancer » même s'il possède son promoteur et peut-être soumis à des régulations post-transcriptionnelles.

Le mode d'action des NAT n'est encore que très partiellement connu. Comme pour les longs ARN non codants, plusieurs mécanismes ont été décrits suggérant une diversité de modes d'action. Il existe cependant plusieurs questions spécifiques à cette catégorie d'ARN comme la possibilité d'une régulation symétrique ou non (l'ARNm peut-il agir en tant qu'ARN antisens de son antisens ?) ou l'importance de la complémentarité de séquence entre les deux transcrits (par opposition avec le fait qu'ils soient transcrits à partir du même domaine chromatinien).

### ***3.2 – Mode d'action : différence et similarité avec les longs ARN non codants***

Les longs ARN antisens constituant un sous-groupe des longs ARN non codants, ils utilisent pour une part les mêmes modes d'action aux niveaux de l'épigénétique, de la transcription et post-transcriptionnel.





## 3.2.1 – Modulation de la chromatine

Parmi les longs ARN non codants permettant de modifier l'état chromatinien, on retrouve un groupe d'ARN qui sont transcrits en antisens. C'est le cas de l'ARN ANRIL qui agira en *cis* et du PTENP1-asRNA $\alpha$  qui agira en *trans*.

### Action en *cis* d'ANRIL

ANRIL ou « Antisense Noncoding RNA in the INK4 Locus » est, comme son nom l'indique, un ARN non codant transcrit sur le brin antisens du locus INK4. Cet ARN régule l'expression des gènes p15/INK4B, p14/ARF et p16/INK4A présents dans le même locus. Le promoteur d'ANRIL se situe entre les gènes p15/INK4B et p14/ARF (Fig. 12).

ANRIL recouvre dans sa totalité le gène suppresseur de tumeur p15/INK4B, d'où l'appellation « p15as » qui a aussi été utilisée. Les trois gènes régulés par ANRIL codent tous des inhibiteurs de kinase cycline-dépendante impliqués dans le contrôle du cycle cellulaire. Ces inhibiteurs bloquent le cycle cellulaire et donc la prolifération. L'expression d'ANRIL favorise donc la tumorigenèse car elle permet de lever l'inhibition de la prolifération. Ainsi la surexpression d'ANRIL est détectée dans certains cancers (Yap *et al*, 2010).

ANRIL est transcrit par l'ARN Polymérase II sous plusieurs isoformes. La majeure partie de ces isoformes est épissée et polyadénylée (Folkersen *et al*, 2009) mais il existe une isoforme non épissée de 34,8 kilobases et également plusieurs isoformes circulaires non polyadénylées. Certaines de ces isoformes ont été montrées comme ayant une expression spécifique de tissu (Folkersen *et al*, 2009 ; Burd *et al*, 2010).

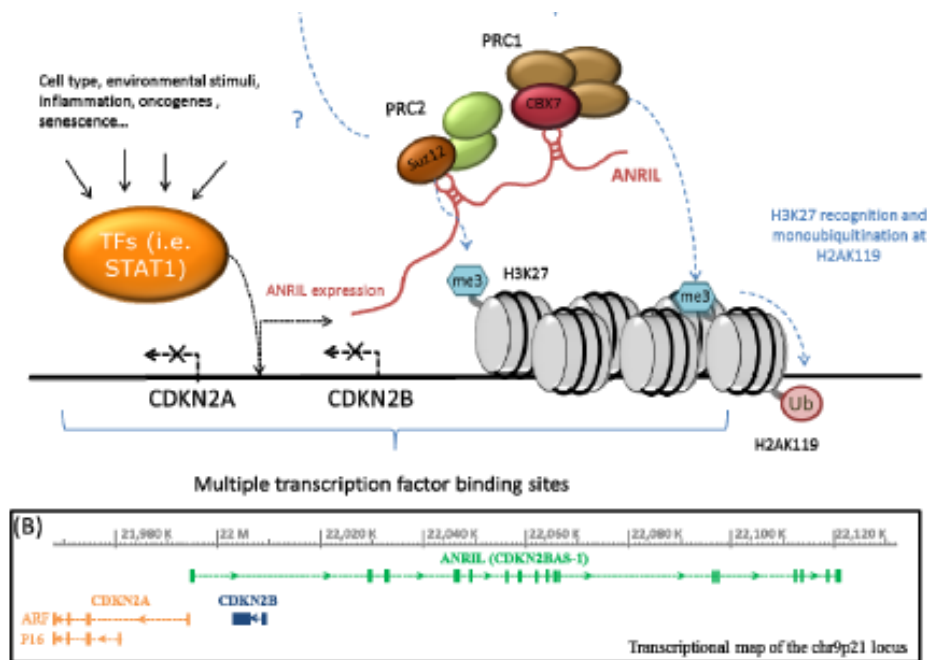
En termes de mode d'action, ANRIL partage avec d'autres longs ARN non codants agissant sur la chromatine le recrutement de complexes modificateurs d'histones. Ainsi il a été montré que plusieurs structures tiges boucles au sein du transcrit ANRIL permettent la liaison au complexe Polycomb PRC2, via CBX7 (Kotake *et al*, 2011 ; Yap *et al*, 2010). En recrutant le complexe PRC2 au sein du locus où il est exprimé, ANRIL permet la pose de la marque inactivatrice de la transcription H3K27me3 et donc le silencing de l'ensemble du locus (Fig. 12).

Bien que son action en *cis* ait été démontrée, des expériences d'expression en *trans* d'ANRIL ont aussi été faites (Yu *et al*, 2008). Ces dernières montrent que même en *trans*, ANRIL est



# Introduction

capable de promouvoir un silencing du gène p15, avec toutefois un effet plus faible que celui en *cis*. Ainsi une transcription locale de l'ARN antisens ne semble pas absolument nécessaire - du moins dans ces expériences. Cela pose aussi la question de la reconnaissance du gène cible, *a priori* via une complémentarité avec les transcrits naissants à moins que ce ne soit avec d'autres transcrits issus du gène comme les « promoter associated RNA » (Schmitz *et al*, 2010).



**Figure 12 : Schéma du mode d'action de l'ARN non codant ANRIL** (Congrains *et al*, 2013).

Le promoteur d'ANRIL est situé entre les gènes p14/ARF/CDKN2A et p15/INK4B/CDKN2B. Il régle l'expression de l'ensemble de son locus qui comprend des gènes impliqués dans l'inhibition de la prolifération. ANRIL lorsqu'il est exprimé permet le recrutement en *cis* des complexes répresseurs Polycomb PRC1 et PRC2 via des structures secondaires au sein de sa séquence. Le recrutement de ces complexes permet la pose des marques H3K27me3 sur les histones du locus et réprime de son expression.

## Action en trans par PTENP1-asRNAα

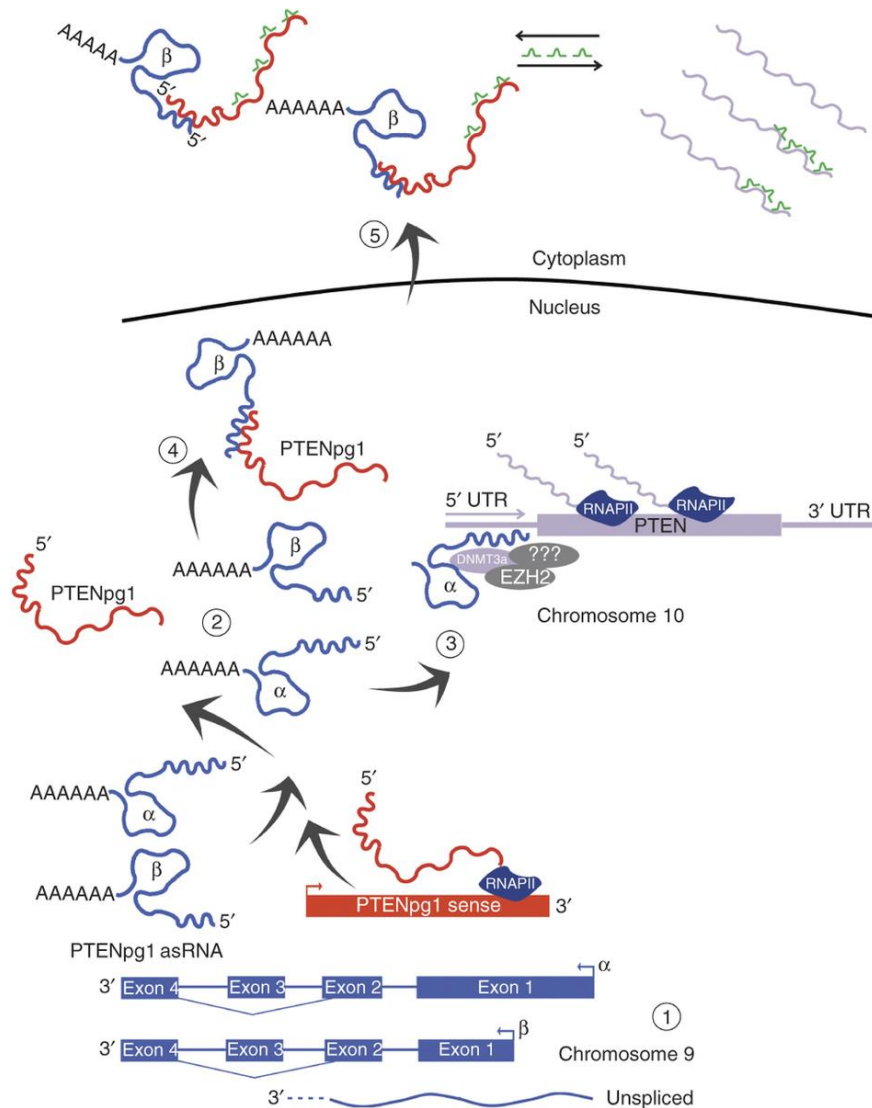
Le cas des ARN antisens agissant sur la chromatine en *trans* est beaucoup plus rare que ceux en *cis*. Le cas le plus connu est celui de l'ARN HOTAIR (déjà présenté précédemment), cependant ce dernier se comporte plus comme un long ARN non codant standard car il n'y a vraisemblablement pas de complémentarité de séquence HOTAIR et ses cibles. Pour retrouver le cas d'un ARN antisens agissant en *trans* sur l'état chromatinien de sa cible par une complémentarité de séquence, il faut se tourner vers les pseudogènes et plus



précisément vers le pseudogène PTENP1. En effet, en plus d'exprimer en « sens » le long ARN non codant correspondant au pseudogène PTENP1, ce locus génère aussi un transcrit antisens, PTENP1-asRNA $\alpha$  (Johnsson *et al*, 2013). PTENP1-asRNA $\alpha$  est l'une des deux isoformes d'antisens produites par le gène PTENP1 qui diffèrent par la position de leur site d'initiation de la transcription (Fig. 13). Cet ARN est transcrit par l'ARN Polymérase II, épissé et polyadénylé. PTENP1-asRNA $\alpha$  régule négativement l'expression du gène PTEN (Johnsson *et al*, 2013). Son mode d'action est standard en recrutant des complexes modificateurs d'histones en interagissant avec la sous unité EZH2 de Polycomb mais aussi avec une méthylase de l'ADN, DNMT3A (Fig. 13). Ainsi en se fixant au niveau du promoteur du gène PTEN, l'ARN antisens permettra le silencing au niveau de celui-ci. L'ARN antisens PTENP1-asRNA $\alpha$  est essentiel à la mise en place du silencing (Johnsson *et al*, 2013). La découverte de cet ARN est récente et la nature exacte de l'interaction entre l'ARN et sa cible n'est pas encore été définie. Il est possible qu'elle soit due à une hybridation avec le transcrit PTEN naissant ou avec des transcrits associés avec le promoteur de PTEN. A noter que dans ce cas, PTENP1-as peut aussi agir en cis sur l'expression de PTENP1 soit par un conflit transcriptionnel soit par hybridation avec les transcrits sens de PTENP1 avec lesquels la complémentarité est évidemment parfaite.

Ainsi dans le cas d'une régulation de la chromatine par un ARN antisens ou NAT, les modes d'action sont les mêmes que ceux des longs ARN non codants standards. L'organisation peut cependant contribuer à la régulation à plusieurs niveaux. A minima le site de transcription antisens étant proche du gène cible, le transcrit naissant peut recruter à proximité de la chromatine cible les activités modificatrices des histones. Alternativement, la formation d'un duplex avec l'ARNm peut soit contribuer à la localisation soit servir de point d'entrée pour un autre type de régulation par exemple une forme d'interférence ARN transcriptionnelle. Une des difficultés conceptuelles de ce mode de régulation est qu'a priori elle peut aussi bien agir sur la transcription du gène cible que de l'antisens. Une expression mutuellement exclusive telle qu'elle est observée dans certains cas comme ANRIL p15 ne découle pas immédiatement des mécanismes envisagés.





**Figure 13 : Schéma du mode d'action des ARNs non codants PTENP1-as** (Johnsson *et al*, 2013).

Les deux isoformes d'ARN PTENP1-as  $\alpha$  et  $\beta$  sont transcrit à partir du pseudogène PTEN sur le chromosome 9. PTENP1-as  $\alpha$  a un rôle répresseur de l'activité du gène PTEN. Ce dernier agit en *trans* et permet le recrutement au niveau du promoteur de PTEN d'éléments régulateurs négatifs comme DNMT3 ou la sous-unité EZH2 du complexe Polycomb. PTENP1-as  $\beta$  a lui un rôle positif sur l'expression de PTENP1. Celui-ci se fixe en 5'UTR de l'ARN non codant PTENP1 et augmente sa stabilité.

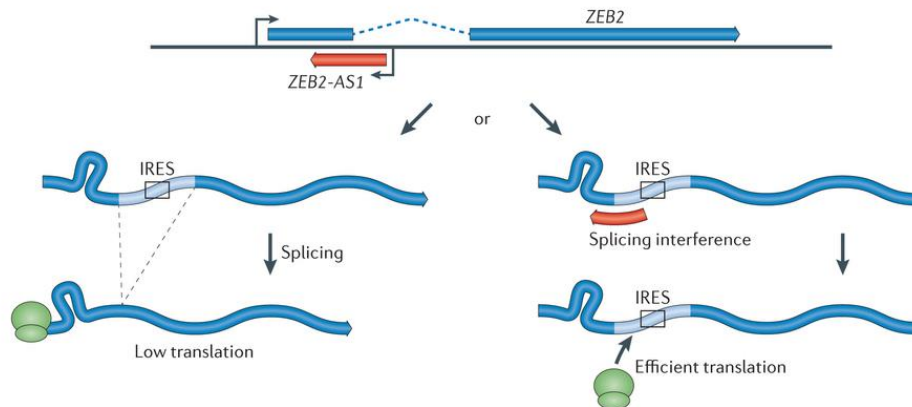
## 3.2.2 – Régulation post-transcriptionnelle

La complémentarité partielle ou complète que peut avoir un ARN antisens avec un gène sens en fait un parfait candidat à la régulation post-transcriptionnelle. En effet, l'ARN antisens va pouvoir former un duplex ARN-ARN avec l'ARN messenger sens qui va pouvoir « interférer » avec l'épissage, la distribution cellulaire, le transport, la stabilité et la traduction de ce dernier. Il existe plusieurs exemples de ce type d'ARN antisens. BACE1-AS, déjà décrit précédemment, en est un qui permet la stabilisation du transcrit sens.



# Introduction

Le cas de l'ARN Zeb2AS est également intéressant. Ce dernier est un transcrit d'environ 1,2 kilobases qui recouvre l'extrémité 5' du gène Zeb2. Zeb2 est un répresseur transcriptionnel de l'E-cadherine et son expression est un élément important de la transition épithélio-mésenchymateuse (EMT) (Neltran *et al*, 2008). La traduction de l'ARN messager Zeb2 nécessite la présence d'un site d'entrée interne du ribosome ou IRES localisé dans le premier intron. Or cet intron est épissé de l'ARN messager mature dans les cellules épithéliales empêchant l'expression de la protéine Zeb2 (Fig. 14). Durant la transition épithélio-mésenchymateuse, l'ARN Zeb2AS est transcrit sous le contrôle du facteur de transcription Snail1 et cet ARN antisens complémentaire du site d'épissage du premier intron empêche la fixation du spliceosome. Ceci permet le maintien de l'intron 1 contenant l'IRES dans l'ARN messager mature et donc sa traduction (Beltran *et al*, 2008).



**Figure 14 : Schéma du mode d'action de l'ARN non codant Zeb2as** (Pelechano et Steinmetz, 2013).

L'ARN Zeb2as est transcrit en antisens depuis le premier intron du gène Zeb2. En absence d'expression de l'ARN Zeb2as, l'intron est épissé privant le transcrit mature d'une traduction efficace. En présence de l'ARN Zeb2as, ce dernier se fixe sur le pré-messager empêchant l'épissage de l'intron 1. L'intron 1 contenant un site d'entrée interne du ribosome le messager mature contenant l'intron 1 peut être efficacement traduit.

Un autre mécanisme susceptible d'être mis en place lors d'une hybridation ARN-ARN est l'interférence par l'ARN. Les ARN antisens grâce à leur complémentarité avec les transcrits sens peuvent en principe constituer une source endogène de siARN. Cela a été démontré chez les mammifères dans l'oocyte de souris (Watanabe *et al*, 2008). Dans cette étude, il est montré que dans le locus Pdzd11/Kif4, où les deux gènes se recouvrent en orientation « head-to-head », on retrouve un grand nombre de petits ARN ciblant cette région. Sur les 135 petits ARN identifiés 117 ciblent différentes séquences de l'exon 1 du gène Kif4. Or c'est exactement l'exon 1 du gène Kif4 qui est recouvert par le transcrit Pdzd11. Il en a été déduit que ces petits ARN interférents résultaient du duplex formé par les ARN des deux transcrits.



Dans la même étude il est observé qu'une mutation de Dicer diminue la quantité de petits ARN ciblant cette région tandis que l'expression des deux gènes est augmentée (Watanabe *et al*, 2008). Ces résultats tendent à confirmer que la transcription antisens permet la création d'une source d'endo-siARN. Néanmoins, certaines données semblent montrer que ces observations ne correspondent qu'à des cas particuliers tissus spécifique aux cellules de la lignée germinale (Werner *et al*, 2014). Dans ce type de mécanisme on peut s'attendre à une régulation similaire sur les gènes sens et antisens.

### 3.2.3 – Conflit transcriptionnel

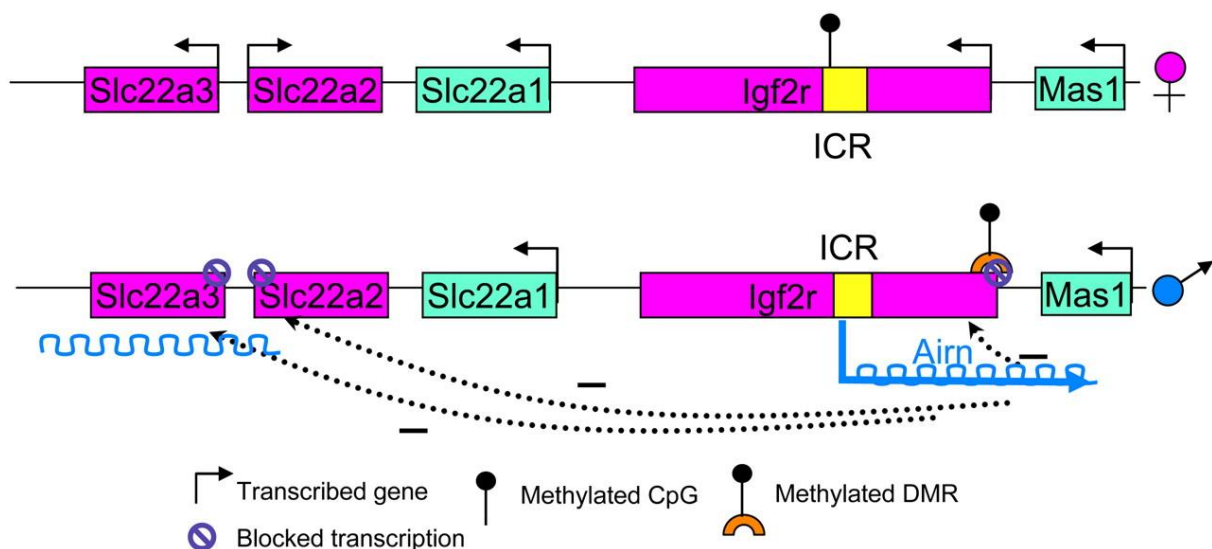
Du fait de leur proximité avec leur gène cible en particulier dans le cas des cis-NAT, un autre mode d'action peut être envisagé : le conflit transcriptionnel. Dans ce type de mécanisme, ce n'est pas l'ARN en lui-même qui va être la source du silencing mais plutôt sa transcription. A titre d'exemple, dans le cas de certaines  $\alpha$ -thalassémies il a été identifié une délétion dans la région 3' du gène LUC7L. Cette délétion fait disparaître le site de polyadénylation de ce gène permettant au transcrit néoformé LUC7L de recouvrir le gène voisin HBA2. Cela a pour conséquence le silencing du gène HBA2, cause de la pathologie (Tufarelli *et al*, 2003). En parallèle de ce silencing on observe la méthylation du promoteur du gène HBA2. Cette méthylation n'est pas une simple conséquence de l'arrêt de la transcription puisque la délétion de la TATA box n'induit pas de méthylation. Elle n'est pas non plus due à une séquence spécifique de l'ARN LUC7L car le remplacement de ce dernier par un autre gène donne le même résultat (Turafelli *et al*, 2003). Il y aurait donc un mécanisme de méthylation de l'ADN *de novo* déclenché par le recouvrement d'un promoteur par une transcription antisens mais ce mécanisme et les acteurs impliqués n'ont pas encore été caractérisés.

Néanmoins, la méthylation n'est pas toujours nécessaire pour le silencing comme c'est le cas pour l'ARN Airn. L'ARN Airn est transcrit à partir d'un promoteur antisens situé dans l'intron 1 du gène de « l'Insulin-like growth factor type 2 receptor » (Igf2r) (Lyle *et al*, 2000) (Fig. 15). Il existe sous plusieurs isoformes, la forme non épissée de 108 kilobases et plusieurs formes épissées de 500 bases à 1,4 kilobases (Seidl *et al*, 2006). Le rôle d'Airn est de réguler l'expression des gènes Igf2r, Slc22a2 et Slc22a3 situés dans le locus soumis à empreinte parentale où il est exprimé (Sleutels *et al*, 2002). L'ARN Airn est exprimé uniquement à partir de l'allèle paternel et va permettre le silencing du locus via le recrutement de l'histone



# Introduction

méthyltransférase G9a (Nagano *et al*, 2008). Néanmoins, cette régulation ne concerne que les gènes *Slc22a2* et *Slc22a3*. En effet, le gène *Igf2r* est mis sous silence par un autre mécanisme. C'est la transcription d'*Airn* qui interférera avec celle du gène *Igf2r*, potentiellement en empêchant la fixation de l'ARN polymérase II sur le promoteur du gène *Igf2r* (Latos *et al*, 2012). Cependant, ce silencing n'est pas durable dans le temps car la transcription de l'ARN *Airn* est nécessaire pour le maintenir. De plus, l'arrêt de la transcription de manière prématurée à l'aide d'un site de polyadénylation qui empêche la transcription jusqu'au promoteur du gène *Igf2r* supprimera le silencing. Ce silencing semble refléter une compétition entre les deux promoteurs car si le promoteur du gène *Igf2r* est remplacé par un promoteur plus fort, le silencing diminue. Dans ce cas de figure, on tend vers un modèle d'un promoteur fort capable d'interférer sur l'activité d'un promoteur plus faible via une transcription antisens. Cela ne fait donc appel ni à une séquence spécifique ni à une complémentarité de séquence mais bien à une transcription antisens qui, en passant sur un promoteur, empêche vraisemblablement la fixation ou déplace les facteurs de transcription ou la machinerie de transcription. A noter que si dans le locus *Igf2r* naturel l'ADN du promoteur *Igf2r* est méthylé, ceci n'est pas toujours le cas dans les différentes constructions où un silencing est observé.



**Figure 15 : Schéma du mode d'action de l'ARN non codant *Airn*** (Bartolomei, 2009).

L'ARN *Airn* est transcrit en antisens du gène *Igf2r* dans le locus soumis à empreinte du même nom. *Airn* permet le silencing des gènes *Slc22a2* et *Slc22a3* par recrutement, via sa structure secondaire, de l'histone méthyltransférase G9a. Le silencing du gène *Igf2r* est lui dépendant de la transcription d'*Airn* et du fait qu'elle recouvre son promoteur.





## II – Les régulations cytoplasmiques en réponse à la présence d'ARN db

Pour l'essentiel, les régulations envisagées pour les ARN non codants interviennent dans le noyau. Or, dans le cas des ARN antisens on peut s'attendre à la formation d'ARN double brin pour lesquels il est connu qu'ils peuvent induire des réponses cellulaires dans le cytoplasme. Pour éviter cette difficulté il a été montré que l'expression d'un ARN double brin dépourvu d'une coiffe et d'une queue polyA (a priori non exportable vers le cytoplasme) pouvait induire un silencing spécifique du gène correspondant (Shinagawa *et al*, 2003). A l'inverse, la transfection dans les cellules d'ARN double brin est connue pour induire l'expression de l'interféron et des gènes de la réponse interféron en déclenchant l'immunité innée de la cellule. En effet, l'accumulation d'ARN et en particulier d'ARN double brin dans la cellule peut être reconnue comme la signature d'une infection virale. Il existe donc dans la cellule des protéines qui vont détecter la présence d'ARN double brin comme signature d'une infection virale. Dans le cas des virus à génome ARN, l'ARN double brin est au minimum un intermédiaire indispensable de la réplication.

### 1 – PKR – Protein Kinase RNA-activated

Parmi les différentes réponses à la présence d'ARN double brin, la voie PKR est une des mieux caractérisées. PKR est une protéine kinase cytoplasmique impliquée dans la réponse de type I qui a un rôle important dans la mise en place de la réponse antivirale et anti-prolifératrice déclenchée par une infection virale (Garcia *et al*, 2006). En absence d'ARN double brin, le domaine kinase de cette protéine s'auto-inhibe. La liaison à une molécule d'ARN double brin induit un changement conformationnel de la protéine libérant le domaine kinase. Plusieurs protéines PKR vont pouvoir se fixer sur une même molécule d'ARN double brin. Cette proximité va faciliter la trans-phosphorylation de celles-ci qui vont dès lors être activées et se dissocier de l'ARN double brin (McKenna *et al*, 2007). La protéine PKR une fois active va pouvoir phosphoryler le facteur d'initiation de la traduction eIF2 $\alpha$  ce qui a pour conséquence le blocage de l'initiation de la traduction et une suppression globale de la synthèse protéique de la cellule (Garcia *et al*, 2006).

La protéine PKR possède deux sites de liaison à l'ARN double brin qui agissent en coopération. La taille minimale des ARN auxquels peut se fixer la protéine PKR est d'environ





# Introduction

---

16 paires de base, néanmoins la dimérisation des protéines PKR sur la même molécule est nécessaire pour leur transactivation. Ceci requiert donc une taille minimale de la molécule d'ARN double brin d'environ 33 paires de bases (Bevilacqua et Cech, 1996 ; Manche *et al*, 1992). Il est notable que cette taille limite de reconnaissance de l'ARN double brin fait que les petits ARN double brin régulateurs (siARN ou miARN) n'activent pas efficacement PKR.

## 2 – OAS – OligoAdenylate Synthase

L'OAS est une enzyme de la famille des Polymérase ARN indépendantes de matrice qui catalyse la liaison de deux molécule d'ATP via une liaison phosphodiester 2'-5' créant les 2'-5' Oligoadenylate (Hovanessian *et al*, 1987 ; Kristiansen *et al*, 2011). Cette enzyme apporte la modification 2'-5' Oligoadenylate sur les ARN double brin qu'elle rencontre dans le cytoplasme. Le mode de fixation de cette enzyme sur la molécule d'ARN n'est pas encore complètement clarifié. Il est proposé que l'OAS soit capable de se fixer à tout type d'acide nucléique par interaction électrostatique, même simple brin, mais que la fixation à l'ARN double brin modifie la conformation de l'enzyme libérant le site actif (Kristiansen *et al*, 2011). L'ajout de la modification 2'-5' oligoadenylate sur l'ARN double brin sert de cofacteur pour l'activation de la ribonucléase RNaseL. La RNaseL dégrade alors les ARN double brin mais également des ARN cellulaires simple brin comme les ARN ribosomiques ou messagers conduisant à une inhibition globale de la synthèse protéique (Floyd-Smith *et al*, 1981 ; Chakrabarti *et al*, 2011).

## 3 – RIG-I – Retinoic acid Inducible Gene I

RIG-I est un autre type de senseur que l'on trouve dans le cytoplasme. Ce dernier possède une architecture complexe et comprend deux domaines CARDs (« Caspase Activation Recruitment Domains ») qui vont pouvoir interagir avec l'adaptateur mitochondrial du signal antiviral (MAVS – « Mitochondrial AntiViral Signaling protein »), un domaine hélicase responsable de l'hydrolyse de l'ATP (dépendant de l'ARN) et un domaine C-terminal de liaison à l'ARN. Cette protéine est capable de reconnaître de longs ARN simple brin porteurs de la modification 5'triphosphate retrouvée chez certains virus (Pichlmair *et al*, 2006) mais également les ARN double brin de petite taille, inférieurs à 2 kilobases (Yoneyama *et al*, 2004). La fixation de la molécule d'ARN par RIG-I permet l'hydrolyse de l'ATP via le domaine

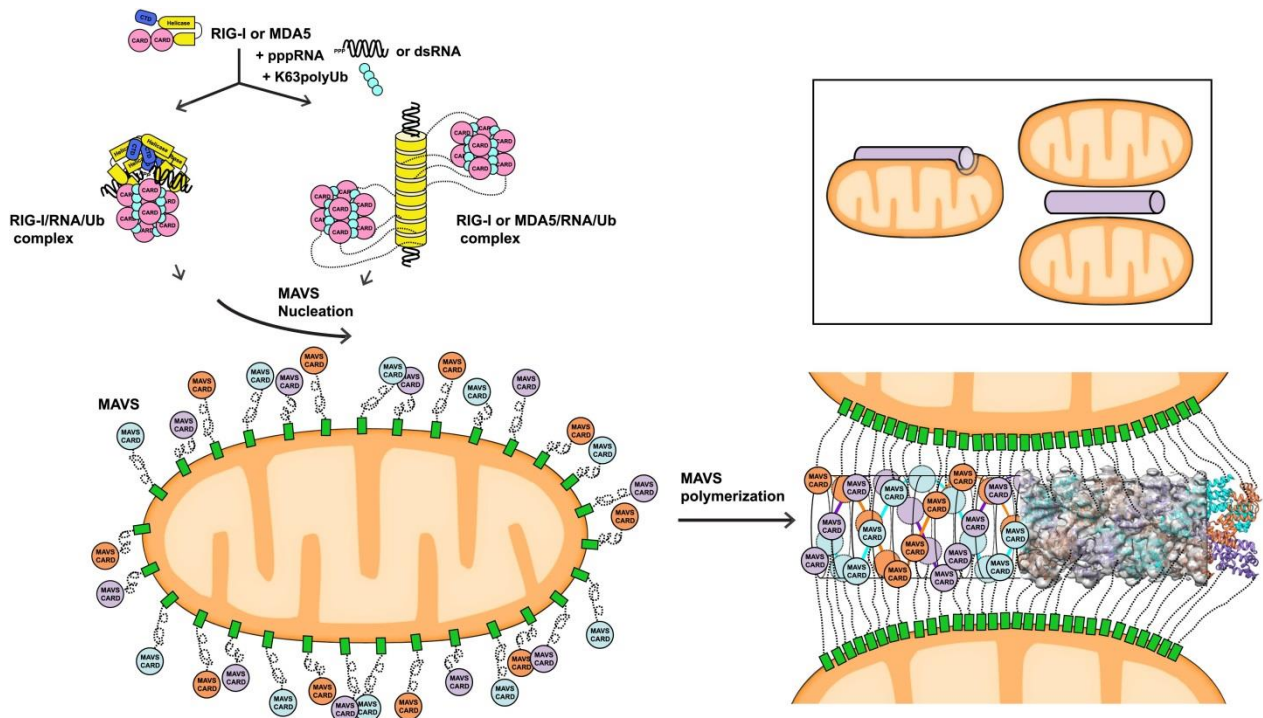


hélicase. Cette hydrolyse de l'ATP rend alors le second domaine CARD sensible à l'ubiquitination par la protéine Trim25 (Gack *et al*, 2007). L'ubiquitination du domaine CARD de RIG-I permet son oligomérisation avec d'autres protéines RIG-I formant alors un filament d'oligomères. C'est ce filament d'oligomères RIG-I qui interagit avec les protéines MAVS en surface des mitochondries (Fig. 16). L'interaction avec MAVS permet la transduction du signal en activant les voies de signalisation comme IRF3/7 et NF- $\kappa$ B résultant en l'expression des gènes de la réponse interféron (Pichlmair *et al*, 2006).

## 4 – MDA5 – Melanoma Differentiation Associated gene 5

MDA5 est également un senseur cytoplasmique de la même famille que la protéine RIG-I. MDA5 possède la même architecture protéique et interagit de la même façon avec l'adaptateur mitochondrial de signal antiviral MAVS. La principale différence est qu'alors que RIG-I permet la reconnaissance des petits ARN double brin (< 2 kilobases), MDA5 interagit avec les ARN double brin de grande taille (> 2 kilobases) (Kato *et al*, 2008). Comme RIG-I, MDA5 va former un filament oligomérique sur l'ARN double brin. Dans ce cas, le filament prend la forme d'anneau empilé tout le long de la molécule d'ARN ou chaque anneau représente un monomère RIG-I (Peisley *et al*, 2011). Mais, contrairement à RIG-I où l'hydrolyse de l'ATP permet l'oligomérisation, ici l'hydrolyse de l'ATP par MDA5 a pour conséquence la dissociation de la protéine de la molécule d'ARN. C'est donc une dynamique instable qui est mise en place. Il est proposé que cette dynamique dissociation-association avec l'ARN double brin permette la discrimination entre les ARN double brin de petite ou grande taille. En effet, si la dissociation est importante, le filament oligomérique de protéine MDA5 sera plus facilement déstabilisé sur une molécule d'ARN de petite taille que sur une de grande taille où le filament, même s'il perd des éléments, restera plus stable (Peisley *et al*, 2011). La formation du filament permet l'interaction entre les domaines CARDS des protéines MDA5 et les protéines MAVS en surface des mitochondries (Fig. 16). Cela permet alors la transduction du signal et induire une réponse antivirale de type interféron.





**Figure 16 : Schéma du mode d'action de RIG-I et MDA5 et leurs interactions avec MAVS** (Xu *et al*, 2014).

Le filament formé par l'interaction entre un ARN double brin de grande taille et MDA5 ou RIG-I permet une accumulation des domaines CARD qui peuvent interagir avec ceux des protéines MAVS en surface des mitochondries. C'est cette interaction qui permet la transduction du signal de la réponse interféron.

## 5 – TLR3 – Toll-like receptor 3

Les Toll-like récepteurs constituent une famille de protéines impliquées dans la réponse immune innée. Ces protéines sont des récepteurs membranaires permettant la reconnaissance des motifs moléculaires associés au pathogène (PAMP – « Pathogen-Associated Molecular Patterns »). Il existe 15 sous-familles de Toll-like récepteurs chez les vertébrés. Elles partagent toutes le même type de structure avec un domaine de liaison au ligand à l'extérieur de la membrane composé de 19 à 25 répétitions en tandem d'un motif riche en leucines qui est lié via un domaine transmembranaire au TIR domaine (TIR – « Toll-like/Interleukin-1 receptor ») impliqué dans l'activation de la voie de signalisation cellulaire. Chaque Toll-like récepteur est spécialisé pour la reconnaissance d'un PAMP donné. En ce qui concerne les acides nucléiques les récepteurs concernés sont TLR3, 7, 8 et 9 (Blasius *et al*, 2010). TLR9 est spécifique de l'ADN contenant des îlots CpG non méthylés (Hemmi *et al*, 2000). TLR7 et TLR8 permettent la reconnaissance des ARN simple brin (Heil *et al*, 2004). Celui qui nous intéresse pour la réponse en présence d'ARN double brin est le récepteur TLR3 (Alexopoulou *et al*, 2001). TLR3 va être, au même titre que RIG-I et MDA5, un senseur

de l'ARN double brin. Mais à la différence de ces deux autres senseurs qui sont cytoplasmiques, TLR3 est présent dans la membrane des endosomes. Ce récepteur permet la détection dans les endosomes de molécules d'ARN double brin d'une taille minimale de 40 paires de bases. La présence d'ARN double brin permet la dimérisation de TLR3 via le domaine externe de liaison à l'ARN double brin et s'accompagne d'une dimérisation du domaine TIR qui permet le recrutement de la protéine TRIF (« TIR-containing adapter-inducing interferon- $\beta$  »). TRIF induit alors l'activation de la voie de signalisation de la réponse antivirale via IRF3/7 et NF- $\kappa$ B.

## 6 – Les protéines ADAR – L'édition ARN

Les protéines ADAR ou « Adenosine deaminase acting on RNA » sont comme leur nom l'indique des enzymes permettant la conversion des adénosines (A) en inosines (I) par désamination de la base. Les protéines ADAR possèdent une région « RNA binding » spécifique de l'ARN double brin ce qui confine l'activité désaminase uniquement à ce contexte (Nishikura et al, 1991). La conversion de l'adénosine en inosine va avoir, sur l'ARN messenger cible, un impact au niveau de sa traduction. En effet, la base inosine est lue par la machinerie de traduction comme une guanosine avec la possibilité d'un changement d'acide aminé et possiblement de l'activité de la protéine codée (Basilio et al, 1962 ; Garrett et al, 2012). Chez l'homme, il existe 3 protéines ADAR. ADAR1 et 2 sont d'expression ubiquitaire tandis que ADAR3 est spécifique du cerveau. ADAR1 et 2 sont les seules à posséder l'activité de déamination (Nishikura et al, 2010). Leur spécificité d'action semble dépendre de leur nombre de « RNA binding Domain », ainsi ADAR2 qui possède 2 « RNA binding Domain » à une spécificité pour certain type de séquence alors qu'ADAR1 qui lui en possède 3 va être moins sélectif (Källman et al, 2003). Une autre différence porte sur leur localisation puisque ADAR1 est présente dans le cytoplasme et le noyau tandis que ADAR2 est principalement nucléaire (Sansam et al, 2003 ; Patterson et al, 1995). ADAR1 est également plus stable que ADAR2 qui est régulée au niveau post-traductionnel par phosphorylation et ubiquitination (Marcucci et al, 2011). Ainsi ADAR1 est logiquement plus fortement exprimé dans les cellules qu'ADAR2 ce qui explique sa découverte plus tardive de celle-ci (O'Connell et al, 1994 ; Kim et al, 1994 ; O'Connell et al, 1997).



# Introduction

---

Par leurs actions, les protéines ADAR sont impliquées dans diverses régulations. Ainsi l'édition ARN permet la formation d'isoformes de protéines et leurs distributions spatiotemporelle au cours du développement (Seeburg et al, 2003 ; Highushi et al, 1993). Les ADAR permettent également par la conversion d'adénosines en inosines de créer ou de faire disparaître des sites d'épissages donnant lieu à de nouveaux variant protéiques (Raitskin et al, 2001). Les protéines ADAR peuvent aussi agir sur les ARN non codant, ainsi il a été montré que les pré-miRNA peuvent être des cibles donnant lieu à des variants de microARN n'ayant plus la même cible (Luciano et al, 2004) mais également sur les ARN des éléments transposables tels que les éléments Alu (Osenberg et al, 2010 ; Bazak et al, 2013). Enfin les protéines ADAR agissent également dans la réponse antivirale en réponse à la présence d'ARN double brin. Comme la conversion en inosine déstabilise les ARN double brin les ADAR jouent potentiellement un rôle important dans la réponse cellulaire aux ARN double brin.



## III – Problématique

L'existence de transcrits antisens a été observée depuis longtemps pour quelques gènes mais ce n'est qu'au cours de ces dernières années qu'il est apparu que, loin d'être une exception, cette situation concernait probablement une majorité de gènes. La perception initiale était que la transcription antisens ne pouvait avoir qu'un rôle inhibiteur pour l'expression du gène « sens » associé. Sans avoir clairement été validés expérimentalement, de nombreux mécanismes inhibiteurs peuvent effectivement être envisagés. Ces mécanismes diffèrent profondément entre eux et par leurs implications pour la régulation du gène cible. La version la plus simple est celle d'un conflit transcriptionnel dans lequel la transcription en antisens empêche la transcription du gène sens que ce soit au niveau de l'élongation (« collision entre polymérase ») ou au niveau de l'initiation (en interférant avec l'assemblage du complexe de transcription sur le promoteur). Dans ce cas, il est clairement indispensable que l'ARN antisens soit exprimé en *cis*. De plus, le mode de régulation est par nature dynamique tant que l'on n'invoque pas l'intervention d'une régulation épigénétique en plus du conflit transcriptionnel ; si la transcription antisens s'arrête le conflit disparaît. Une autre spécificité de ce mode de régulation est que l'expression tend à être mutuellement exclusive puisque la transcription sens peut aussi inhiber la transcription antisens. On peut donc supposer que l'unité la plus transcrite prend le dessus sur l'autre. Une deuxième classe de mécanismes implique l'interaction entre les transcrits sens et antisens et l'induction d'une réponse par l'ARN double brin. Typiquement une interférence par l'ARN au niveau transcriptionnel rentrerait dans cette catégorie. Toutefois l'efficacité de ce type de régulation, en particulier chez les mammifères, est encore un sujet de débat. Potentiellement ce type de mécanisme pourrait avoir lieu si l'ARN antisens est exprimé en *trans* à partir d'un autre locus et le silencing pourrait lui-même agir sur d'autres loci dès lors qu'ils partagent des séquences avec l'ARN double brin. L'existence d'une telle régulation a été proposée par le groupe de N. Proudfoot (Gullerova *et al*, 2012). Toujours liés à la formation d'ARN double brin d'autres mécanismes sont envisageables qui pourraient être indépendants de la séquence. C'est bien évidemment le cas pour les réponses cytoplasmiques (PKR, RNaseL, détection de l'ARN double brin et induction de l'expression de l'interféron ...), mais rien ne permet d'exclure qu'une réponse locale puisse être en place au niveau du site de transcription dans le noyau. A priori ce type de réponse devrait agir aussi



# Introduction

---

bien sur les deux unités de transcription. Ces trois mécanismes bien que très différents partagent le fait qu'ils ne dépendent pas de la nature des séquences transcrites et donc devraient être mis en place systématiquement dès lors que certaines conditions sont satisfaites (organisation des deux unités de transcription, niveau d'expression des deux transcrits et quantité d'ARN double brin formé, taille de la région double brin...). A l'inverse, une majorité des mécanismes décrits pour l'action des longs ARN non codants nécessite la présence sur le transcrit de motifs spécifiques pour permettre le recrutement de complexes régulateurs. Dans ce cas, la formation d'un duplex d'ARN dans le noyau ne semble pas nécessaire voire même pourrait entrer en conflit avec l'accessibilité des sites de recrutement. D'ailleurs, si le cas le plus fréquent est une régulation en *cis*, des exemples de régulation en *trans* ont été décrits dans lesquels il n'existe pas de complémentarité notable entre l'ARN régulateur et sa cible. Dans ce type de modèle, la régulation de l'expression passe par des modifications de la chromatine et est donc susceptible de se propager au-delà de la région où le transcrit antisens naissant recrute le(s) complexe(s) modificateurs. En fonction de la présence de barrières ou d'isolateurs chromatinien, le silencing induit par l'antisens peut donc se propager jusqu'aux deux promoteurs conduisant au silencing des deux gènes. En fonction de la stabilité de l'organisation chromatinienne induite, une telle situation ne créerait pas nécessairement une régulation inefficace, l'expression de l'antisens servant d'évènement initiateur du silencing. De fait, si on peut envisager des moyens de protéger l'expression de l'ARN antisens, en particulier si les promoteurs sens et antisens sont très éloignés l'un de l'autre sur le génome, cette classe de mécanismes n'est pas intrinsèquement asymétrique.

Contrairement à ce qui était attendu, les études sur l'ensemble du transcriptome n'ont pas mis en évidence de corrélation négative entre l'expression des ARN sens et antisens sur l'ensemble du génome. Il existe même de nombreux exemples de co-expression des transcrits sens et antisens. Cette observation est probablement à relier avec le fait que la présence d'une unité de transcription antisens pourrait aussi être activatrice puisqu'en l'absence d'un isolateur transcriptionnel, l'activation du promoteur antisens est a priori susceptible d'augmenter la transcription depuis le promoteur sens. Un tel mécanisme est un simple reflet du phénomène fréquemment observé de transactivation entre promoteurs, mais sa réalité effective dépend directement du poids des mécanismes d'inhibition



mentionnés plus haut.

L'objectif de ce projet est d'évaluer l'impact sur l'expression d'un gène cible des différentes classes de mécanismes d'action des ARN antisens qui ont été évoquées. La démarche est d'augmenter progressivement le nombre de mécanismes susceptibles d'intervenir en partant des modèles les plus simples possibles. Ainsi, à l'inverse des études sur des loci cellulaires qui cherchent à mettre en évidence au cas par cas quel mécanisme est impliqué, il devrait être possible d'évaluer plus globalement l'importance de ces différents mécanismes en fonction du contexte (organisation et expression des gènes sens et antisens). La démarche expérimentale repose sur le suivi par cytométrie de l'expression de constructions indicatrices comprenant les régions codantes de protéine fluorescentes. Outre son caractère quantitatif (ou semi-quantitatif comme il apparaîtra dans les résultats) un avantage important de la cytométrie est d'analyser l'expression au niveau de cellules individuelles et non pas en valeur moyenne sur une population. Ainsi dans l'hypothèse où un silencing ne se mettrait en place que dans une sous-population de cellules, ceci sera parfaitement analysable en cytométrie même si l'impact sur l'expression moyenne n'est pas visible. Toujours dans le but de partir de la situation la plus simple possible, une première partie du projet porte sur l'expression en transitoire des constructions indicatrices. Si l'on peut s'attendre à ce que la chromatinisation des plasmides ne soit que partielle sur cette échelle de temps (typiquement 48 h), ceci ne devrait pas empêcher un conflit transcriptionnel tout comme certains effets liés à la formation d'ARN double brin. A l'opposé, les effets dus au recrutement de complexes modificateurs de la chromatine devraient être nettement plus marqués dans une situation de clone stable et nécessiter la présence de motifs spécifiques de recrutement dans la construction antisens.





## MATERIELS ET METHODES

### I – Cellules

#### 1 – La lignée cellulaire

L'intégralité des expériences discutées dans cette étude a été faite dans la lignée cellulaire HEK 293 Tet-on. La lignée HEK 293, établie à partir de cellules de rein embryonnaire humain au début des années 1970, est fréquemment utilisée comme système modèle de cellules en culture du fait de sa croissance rapide et de sa très grande efficacité de transfection. Ces cellules sont efficacement transfectables par presque toutes les techniques existantes et tout particulièrement par la méthode au précipité de phosphate de calcium. La lignée HEK 293 Tet-on est une lignée dérivée de HEK 293 commercialisée par Clontech. Il s'agit d'un transfectant stable pour le gène rtTA obtenue par sélection avec la généticine (G418). Le gène rtTA code une protéine de fusion entre une version mutée du répresseur de la tétracycline et le domaine transactivateur de VP16. Ainsi contrairement à la version sauvage du répresseur, la fixation du répresseur muté sur son site au niveau de l'ADN va dépendre de la présence de tétracycline. L'addition de tétracycline ou préférentiellement de son dérivé, la doxycycline, dans le milieu de culture va entraîner la fixation du transactivateur sur ses séquences cibles et induire l'activité des séquences promotrices correspondantes. L'induction est progressive au cours du temps et est fonction de la dose de doxycycline. En pratique, la dose utilisée (1 µg/mL) assure une réponse maximale après 24h d'induction. Une des limites de l'utilisation des cellules HEK 293 pour nos expériences est leur faible adhérence au plastique, ce qui oblige à attendre 48h après ensemencement pour effectuer une transfection. Néanmoins, cette caractéristique devient un avantage pour l'analyse en cytométrie en flux puisqu'elle permet d'obtenir une suspension cellulaire assez facilement par ajout d'un chélateur de calcium.



## 2 – Culture cellulaire

La lignée HEK 293 Tet-on et ses dérivés avec constructions intégrées sont cultivées en étuve humidifiée à 37° et à 5% de CO<sub>2</sub>. Pour le maintien des cellules en culture, les cellules sont divisées lorsqu'elles atteignent 80% de confluence par dissociation avec une solution trypsine-EDTA et réensemencées à des dilutions variant de 1/6 à 1/20 en fonction de la durée prévue de culture. Les cellules sont cultivées dans du DMEM (avec 4,5 g/L glucose et sans pyruvate, Gibco) supplémenté par 10% de sérum de veau fœtal (SVF, Gibco), 100 unités/mL de pénicilline et 100 µg/mL de streptomycine.

## 3 – Transfection

L'ensemble des transfections réalisées au cours de ce travail ont été faites par la méthode au phosphate de calcium. Cette méthode a été mise au point par S. Bachetti et F.L. Graham en 1977. Elle repose sur le mélange d'une solution saline tamponnée par l'HEPES doublement concentrée (HEPES-buffered saline solution ou HBS) contenant des ions phosphates et une solution de chlorure de calcium contenant les acides nucléiques à transfecter (petits ARNs et/ou plasmides). Lorsque ces deux solutions sont mélangées, il se forme un fin précipité de phosphate de calcium dans lequel sont intégrés les acides nucléiques présents en solution. Le mélange contenant les précipités est alors ajouté aux cellules à transfecter (habituellement une culture cellulaire monocouche). C'est par un mécanisme non entièrement compris que les cellules vont incorporer le précipité contenant les acides nucléiques (les ions Ca<sup>2+</sup> masquent la polarité négative de l'ADN lui permettant d'entrer dans le cytoplasme de la cellule). Dans les cellules HEK 293, la technique au phosphate de calcium est très efficace et permet la transfection du plasmide d'intérêt dans pratiquement toutes les cellules. En effet, lors de contrôles effectués avec un ADN conjugué à un fluorochrome, une augmentation de la fluorescence peut être détectée dans toutes les cellules ou presque (> 95 % des cellules). Par contre, le fait que les cellules aient reçu de l'ADN ne conduit pas nécessairement à une expression facilement détectable des constructions transfectées. Ainsi, l'efficacité de transfection déterminée par le pourcentage de cellules dont la fluorescence est supérieure à l'autofluorescence varie de 10 à 30 % en fonction des expériences. Ces cellules sont celles pour lesquelles l'analyse en cytométrie sera la plus facile



à réaliser, même si une augmentation de la fluorescence est aussi détectable pour les autres cellules suggérant qu'il y a bien une faible expression des constructions transfectées dans celles-ci. Une caractéristique essentielle pour nos expériences du modèle HEK 293/transfection au phosphate de calcium est la très grande efficacité de co-transfection avec deux ou trois plasmides et un ou deux ARN interférents. Dans des expériences précédemment réalisées dans l'équipe, cette efficacité de co-transfection n'a pas été observée avec d'autres réactifs de transfection comme la lipofectamine.

Pour les transfections au phosphate de calcium, les cellules sontensemencées 48h auparavant à  $7.10^5$  cellules/ boîte  $\phi$  100mm,  $4.10^5$  cellules/boîte  $\phi$  60mm ou  $1.10^5$  cellules/plaque de 12 puits. Le milieu de culture est renouvelé au minimum 2 heures avant la transfection.

Pour les plaques 12 puits, format des transfections réalisées en routine, les plasmides et petits ARN sont dilués dans une solution de Tris/EDTA diluée 10 fois (concentration finale 1mM Tris pH 7,8 et 0,1mM d'EDTA). Pour chaque échantillon est préparée dans un tube Eppendorf 50  $\mu$ L une solution A contenant les acides nucléiques et 6 $\mu$ L de  $\text{CaCl}_2$  2M, qui est mélangée par vortexage. Par-dessus la solution A est ajouté doucement un volume de HBS 2X (50 $\mu$ L) qui apporte les ions phosphates. La précipitation est déclenchée en initiant le mélange des deux solutions en insufflant avec une micropipette des bulles d'air au fond du tube (quatre fois). Après 20 minutes à température ambiante, le précipité est ajouté sur les cellules (100 $\mu$ L/puits). 24h après la transfection, le milieu de culture est renouvelé avec ou sans doxycycline selon les constructions transfectées. L'analyse par cytométrie en flux se fait le plus souvent 48h après la transfection.

Solution utilisée pour la transfection :

- $\text{CaCl}_2$  2M
- HBS 2X (pH ajusté à 7,05 avec une solution de NaOH)
  - HEPES 50mM
  - NaCl 280mM
  - KCl 10mM
  - $\text{Na}_2\text{PO}_4$  1,5mM



## 3 – Construction des clones stables HEK 293 Tet-on

Pour l'isolement de clones stables, les cellules de la lignée HEK 293 Tet-on sont co-transfectées par la méthode au phosphate de calcium avec le plasmide d'intérêt et un plasmide de résistance à la puromycine « pBabe puro ». Le ratio molaire plasmide indicateur / pBabe puro est de 4 : 1 pour favoriser la probabilité que les clones résistants aient intégré le plasmide indicateur. La transfection est faite dans des boîtes 100 mm avec des cellules ensemencées 48h auparavant à  $7.10^5$  cellules/boîte. Le milieu de culture est renouvelé 24h après la transfection et la mise en place de sélection est faite à 48h par ajout de  $1\mu\text{g/mL}$  de puromycine dans le milieu de culture. Les cellules sont cultivées sous pression de sélection pendant 7 à 10 jours. La sélection est par la suite enlevée pour que les clones puromycine résistants puissent proliférer plus activement. Les clones visibles obtenus ( $\phi$  de plus 2mm) sont transférés dans une plaque 24 puits dans un milieu sans sélection pendant 3 à 4 jours. Puis la puromycine est à nouveau ajoutée pour vérifier la résistance des clones sélectionnés. Les clones sont ensuite analysés par cytométrie en flux afin de déterminer le profil d'expression de la construction intégrée. Les clones ayant un profil d'expression satisfaisant sont mis en culture puis congelés à  $-150^\circ\text{C}$ .

## II – Cytométrie en flux

Les constructions indicatrices utilisées permettent l'expression de protéines fluorescentes (GFP, DsRed et Nir). La cytométrie en flux permet une mesure quantitative de l'expression de ces derniers au niveau de chaque cellule. De plus, comme il s'agit de protéines ayant des fluorescences intrinsèques, il n'est pas nécessaire de perméabiliser ou de fixer les cellules. Pour disposer d'un nombre suffisant de données exploitables, la lecture est faite sur 50 000 cellules d'où le choix d'utiliser des plaques 12 puits ensemencées à  $1.10^5$  cellules.

### 1 – Préparation des échantillons

Dans les plaques 12 puits, le milieu est aspiré puis sont ajoutés  $200\mu\text{L}$ /puits de PBS 1X contenant 20% de tampon de dissociation cellulaire (Gibco). Il s'agit d'un milieu tamponné sans  $\text{Ca}^{2+}$  et contenant une quantité non spécifiée d'EDTA. Les cellules sont alors dissociées



en pipetant vigoureusement la solution avec une pipette P1000, puis transférées dans un tube sur glace (Tube polystyrène 5mL pour analyse avec le FACS Calibur de Becton Dickinson ou Tube Eppendorf 1,5mL pour analyse avec le GUAVA de Millipore). Les échantillons peuvent être gardés dans la glace et à l'abri de la lumière pendant quelques heures sans que les résultats en soient affectés comme on peut le vérifier en analysant l'échantillon à différents temps de cinétique.

## 2 – Protocole d'analyse

### 2.1 – Analyse en deux couleurs : GFP et DsRed

#### Acquisition

Pour analyser l'expression de la GFP et de la DsRed à partir des constructions indicatrices, il a été utilisé l'excitation par le laser bleu 488nm présent sur les deux cytomètres en flux utilisées durant notre travail ainsi que les premiers détecteurs de fluorescence FL1 (533  $\pm$ 15nm) et FL2 (582  $\pm$ 20nm) sur le FACS Calibur (Becton Dickinson) et les détecteurs Green-B (525  $\pm$ 30nm) et Yellow-B (583  $\pm$ 26nm) sur le GUAVA (Millipore). Cette combinaison est parfaitement adaptée à la détection de la GFP (excitation max : 488nm ; émission max : 509nm) mais cependant un peu moins pour la DsRed (excitation max : 557nm ; émission max : 579nm) puisque l'excitation par le laser à 488nm ne permet que 50% de l'excitation maximale. La lecture est faite sans compensation pour les passages d'une fluorescence sur une autre. La compensation est faite *a posteriori* au cours de l'analyse des données ce qui permet de tester la robustesse des conclusions vis à vis du choix des valeurs de compensation (voir plus bas). Pour chaque échantillon et quel que soit le cytomètre en flux utilisé, sont analysées au minimum 50 000 cellules avec un seuil sur la diffusion vers l'avant afin d'éliminer les débris les plus petits dès l'acquisition. L'acquisition sur le FACS Calibur (Becton Dickinson) pose plusieurs difficultés car l'analyse ne peut se faire que sur 4 ordres de grandeur, ce qui, compte tenu de la saturation des détecteurs observée aux signaux les plus forts, se réduit à 3,5 décades exploitables. Or la gamme de signaux que nous observons entre les cellules non transfectées (autofluorescence) et les plus forts expresseurs est proche de 5 décades. Aussi pour l'acquisition des données sur FACS Calibur, deux lectures sont effectuées avec des réglages de sensibilité des détecteurs de fluorescences différents afin de



couvrir l'ensemble des niveaux d'expression. Le cytomètre en flux GUAVA permet une quantification des signaux sur 5 décades ce qui permet une acquisition en une seule lecture de l'ensemble des niveaux d'expression. Le GUAVA n'utilise pas le même système de fluide que le FACS Calibur. Son système est plus simple et repose sur le passage des cellules dans un capillaire situé devant le faisceau laser, la circulation des cellules étant assurée par une seringue motorisée. En contrepartie d'une fluidique beaucoup plus simple, le profil de diffusion est moins propre du fait de la présence du capillaire et d'une moins bonne séparation entre cellules. Néanmoins, malgré ces différences techniques, le GUAVA est plus facile d'utilisation et suffisant pour nos analyses. C'est pourquoi il est utilisé pour la majorité des résultats acquis dans cette étude.

## Acquisition

Pour l'analyse des données acquises deux logiciels ont été utilisés. Le logiciel Weasel (Walter and Eliza hall Institute, Melbourne) pour les données obtenues par le FACS Calibur et le logiciel Incyte GuavaSoft pour les données du GUAVA. Dans les deux cas, l'analyse est menée de la même manière avec tout d'abord une sélection sur la diffusion qui permet d'éliminer les cellules mortes, les débris et les agrégats. En général au moins 80% des événements sont retenus et les valeurs finales ne varient pas fortement en fonction du choix de la sélection sur le profil de diffusion. Dans l'analyse de base, l'expression moyenne est obtenue à partir des valeurs moyennes globales de la fluorescence obtenues dans les canaux GREEN-B et YELLOW-B (FL1 et FL2 pour le FACS Calibur). En fonction du contexte, d'autres analyses ont été effectuées avec des sélections ou seuils pour exclure certaines sous-populations.

## Compensation

Le réglage de compensation est défini en utilisant des cellules transfectées uniquement avec le vecteur pCMV GFP ou pCMV DsRed afin de mesurer le passage d'une fluorescence dans l'autre. Le pourcentage de passage d'une fluorescence dans l'autre va dépendre du cytomètre utilisé (du fait des paramètres physiques comme les longueurs d'onde sélectionnées par les filtres) mais aussi du réglage de la sensibilité des détecteurs). Dans nos expériences pour le FACS Calibur, le passage de signal GFP dans le canal DsRed est d'environ 20% tandis que l'inverse, le passage de la DsRed dans le canal GFP est de 2%. Pour le GUAVA,



le passage de la GFP dans le canal DsRed est de 7,4% alors que l'inverse est de 2,2%.

Une difficulté majeure rencontrée dans nos expériences vient de l'impossibilité d'effectuer une compensation qui soit efficace pour tous les niveaux d'expression. Dans le cas du FACS Calibur le manque de linéarité des détecteurs est flagrant et seule une compensation « locale » est réellement possible (sur une décade). Les performances du GUAVA, appareil plus récent, sont un peu meilleures et permettent une compensation régionale (sur trois décades). Néanmoins, il n'est pas possible d'avoir une compensation qui permette de traiter correctement l'ensemble du spectre d'expression observé dans nos cellules. Ce défaut des cytomètres n'a pas encore été complètement corrigé même dans les appareils les plus récents du fait qu'il n'ont pas été conçus pour faire des analyses quantitatives mais qualitatives afin d'identifier des différences d'expression sans que la valeur de cette différence constitue un enjeu. De plus, lors des expériences de transfection transitoire nous observons une distribution de fluorescence sur cinq décades ce qui représente probablement une situation extrême par rapport aux niveaux d'expression des gènes cellulaires.

## *2.2 – Analyse en trois couleurs : GFP, DsRed et Nir*

Pour la présence d'un contrôle externe au silencing ou un contrôle d'expression en transitoire, la construction pCMV Nir a été utilisée. Les cellules sont alors analysées au cytomètre de flux en utilisant le canal Red-R ( $661 \pm 15\text{nm}$ ) sur le GUAVA ou FL4 ( $675 \pm 12\text{nm}$ ) sur le FACS Calibur. L'excitation de la protéine fluorescente Nir (excitation max : 605nm ; émission max : 670nm) par le laser 642nm présent sur les deux cytomètres est peu efficace. Le passage de Nir ne se fait que dans le canal DsRed et reste modéré. Les valeurs de compensation sont de 6% de passage de Nir dans le canal DsRed sur le FACS Calibur et de 0.2% sur le GUAVA.

## **III– Biologie moléculaire**

### **1– Les plasmides**

Tous les plasmides rapporteurs sont produits dans la souche bactérienne Library Efficiency DH5 $\alpha$  (Invitrogen) et purifié avec le Kit Nucleobond Xtra Midi de Macherey-Nagel.



## 1.1 – Les constructions pBiFluo et dérivés

Le vecteur pBiFluo a été élaboré à partir du vecteur pBiFluo-Glo (Fig. M1a) initialement construit par Maria Antonietta Buccheri. Celui-ci avait été obtenu par l'intégration dans un plasmide contenant un promoteur bidirectionnel régulé par la tétracycline (pBi, Clontech), des séquences codant la DsRed Express (Clontech) et d'un gène de fusion entre l'EGFP et la  $\beta$ -Globine de lapin (contenant les deux introns). La digestion du plasmide pBiFluo-Glo par BglII a permis de libérer la région  $\beta$ -globine et c'est la re-circularisation du plasmide sur lui-même qui permet d'obtenir le plasmide pBiFluo (Fig.M1b).

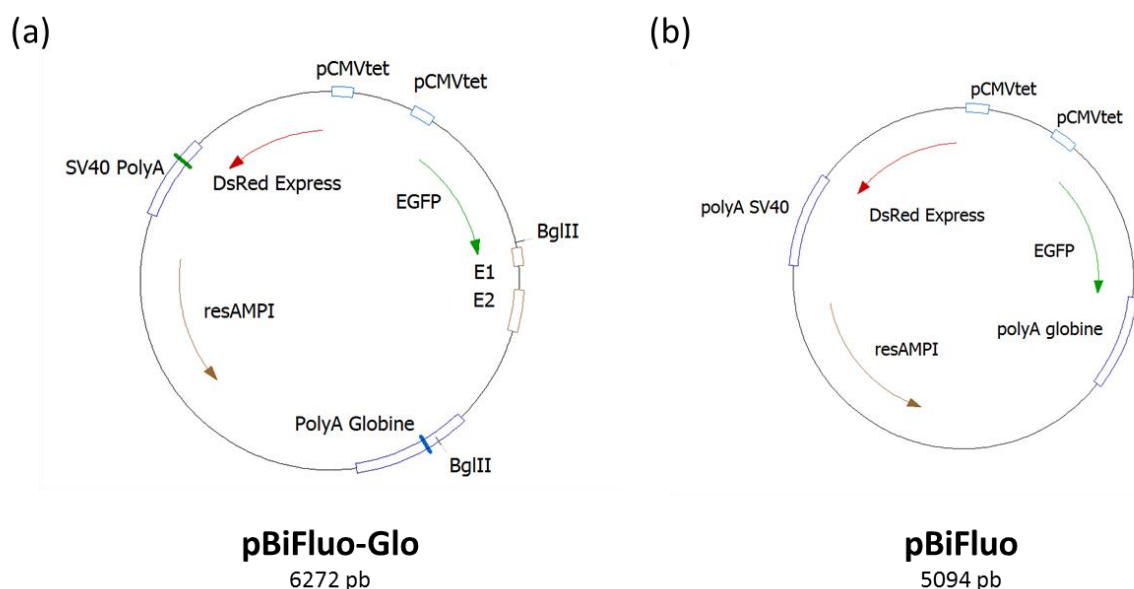


Figure M1 : Carte des plasmides pBiFluo-Glo (a) et pBiFluo (b)

Le site de l'enzyme de restriction BglII utilisé pour déléter la région  $\beta$ -Globine et créer pBiFluo est affiché sur la carte. Carte obtenue via le logiciel Clone manager (Sci Ed Central).

L'EGFP diffère de la protéine initiale de méduse (*Aequorea victoria*) par deux mutations (F64L et S65T) qui améliorent son repliement et augmente le rendement de fluorescence tout en déplaçant le spectre d'excitation vers des valeurs proches de celles des instruments habituellement utilisés (de 395nm à 488nm). La DsRed est une protéine tétramérique issue d'un corail (*Discosoma*) capable de fluorescer dans le rouge. Comme pour la GFP sauvage, le repliement de la DsRed est lent et peu efficace ce qui peut être en partie contourné par l'introduction de mutations. La DsRed express est un variant qui a une maturation plus





rapide dans les cellules de mammifère. Néanmoins, son expression sous forme fonctionnelle reste plus lente que celle de la GFP.

Le promoteur bidirectionnel qui contrôle l'expression des plasmides de la série pBiFluo ainsi construits contient deux promoteurs pTRE identiques régulés par la tétracycline (ou la doxycycline). Chaque promoteur pTRE consiste en une séquence promotrice CMV minimal ( $P_{minCMV}$ ) et une séquence régulatrice TRE (Tet Response Element) en amont qui régule l'activité du premier (Fig. M2). La séquence TRE induit la transcription du gène contrôlé par le promoteur pTRE (dans notre cas EGFP ou DsRed) quand le transactivateur protéique s'y fixe. Dans les lignées cellulaires Tet-On, qui expriment constitutivement le transactivateur rtTA, l'expression des plasmides pBiFluo et dérivés est induite par la doxycycline (DOX) car sa fixation sur rtTA permet à ce dernier d'interagir avec la séquence TRE, comme présenté dans la Figure M2.

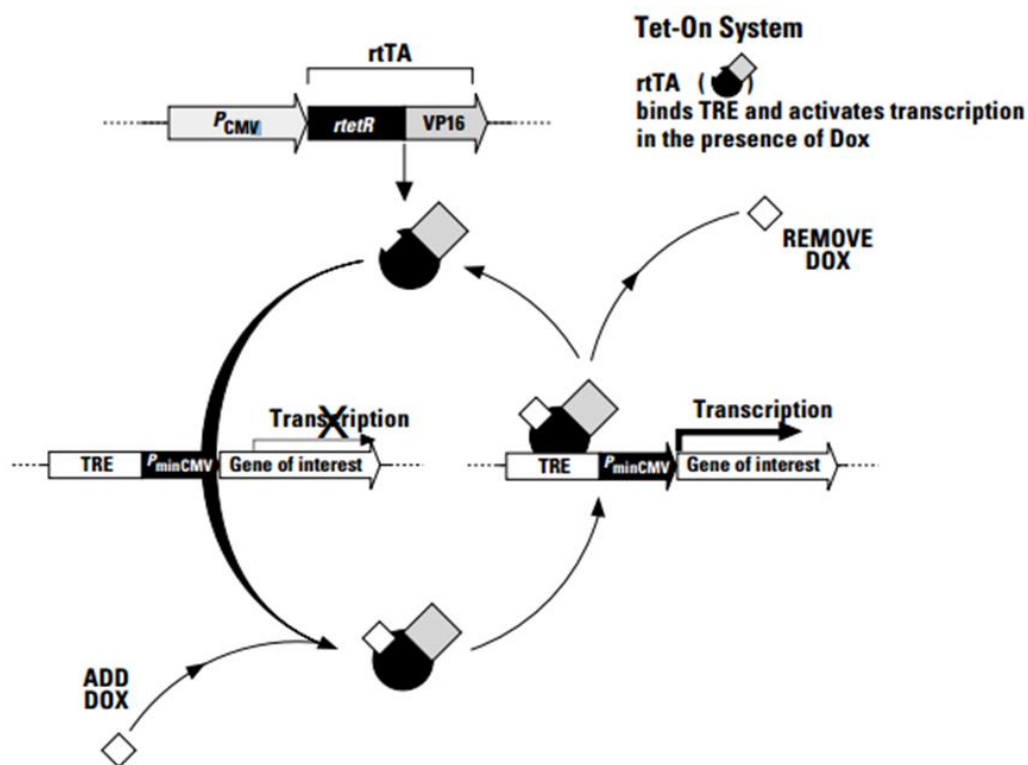
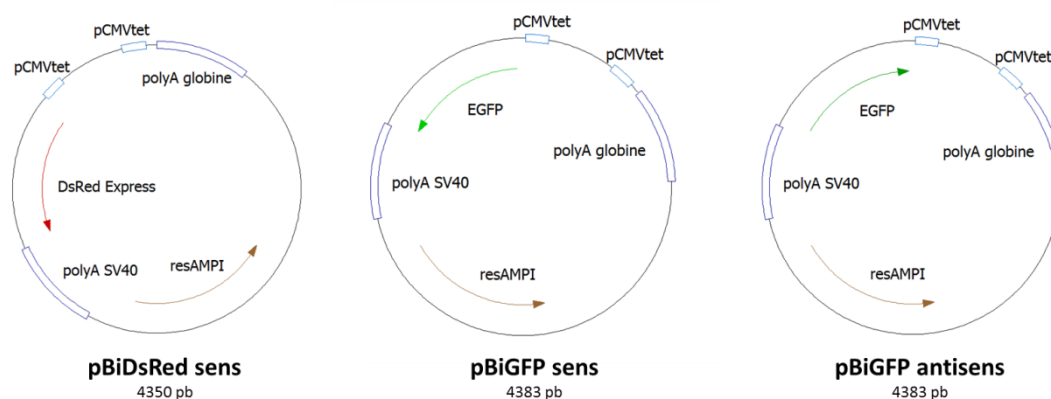


Figure M2 : Schéma du système d'expression Tet-on

De cette construction pBiFluo dérive plusieurs autres constructions. En premier lieu des constructions monochromes issues de l'excision de la GFP donnant ainsi pBiRed. Mais aussi via l'excision de la séquence DsRed et son remplacement par les séquences GFP, Nir ou DsRed en orientation antisens pBiRed antisens, pBiNir sens, pBiNir antisens et pBiGFP antisens (Fig. M3). C'est ainsi qu'ont été créées les constructions antisens sous dépendance du promoteur inducible à la tétracycline.

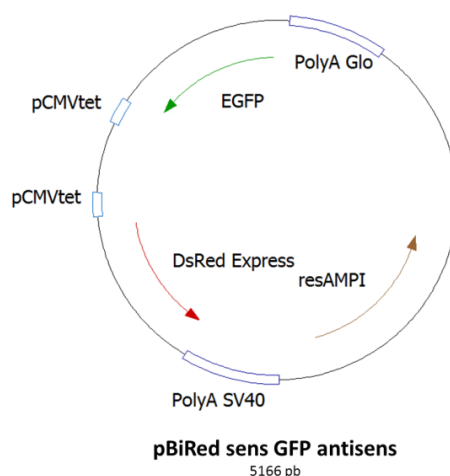




**Figure M3 : Carte des constructions monochromes et antisens dérivées du pBiFluo**

Carte uniquement de 3 constructions issue de pBiFluo, pBiRed sens, pBiGFP sens et pBiGFP antisens

C'est également à partir de la construction pBiFluo qu'a été réalisée la construction pBiRed sens GFP antisens utilisée en particulier pour les clones stables pour suivre l'expression de la construction antisens. Cette dernière est le produit de l'excision de la GFP de pBiFluo puis la réinsertion de cette même GFP mais cette fois-ci en orientation antisens par rapport au promoteur (Fig. M4).

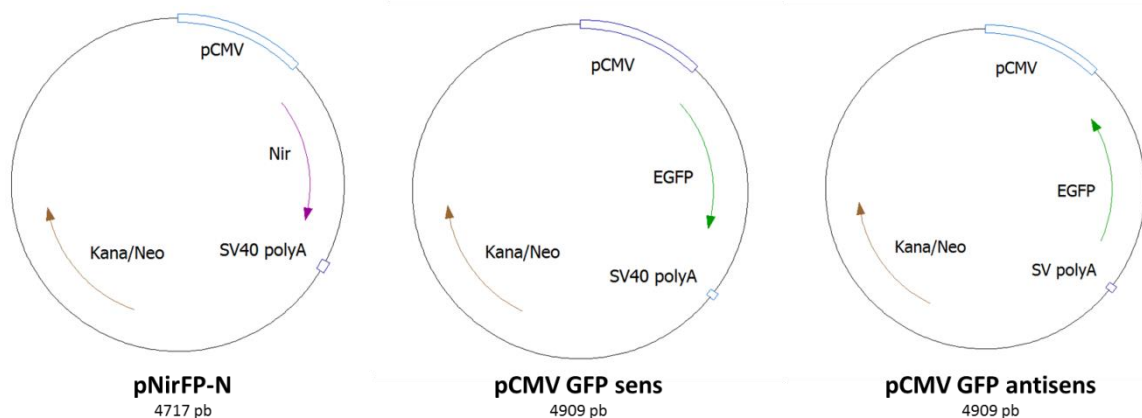


**Figure M4 : Carte des constructions pBiRed sens GFP antisens**

## 1.2 – Les constructions indicatrices CMV

Contrairement aux constructions pBi, les constructions CMV vont toutes être monochromes. En suivant la même méthode, les constructions sous dépendance du promoteur constitutif CMV ont été réalisées à partir du plasmide commercial pNirFP-N (Evrogen), aussi appelé pCMV Nir sens. Les variants GFP, DsRed et antisens pour les différents gènes de protéines fluorescentes ont été créés en déléant la séquence codant la protéine Nir et en y insérant les séquences en sens ou antisens (Fig. M5).



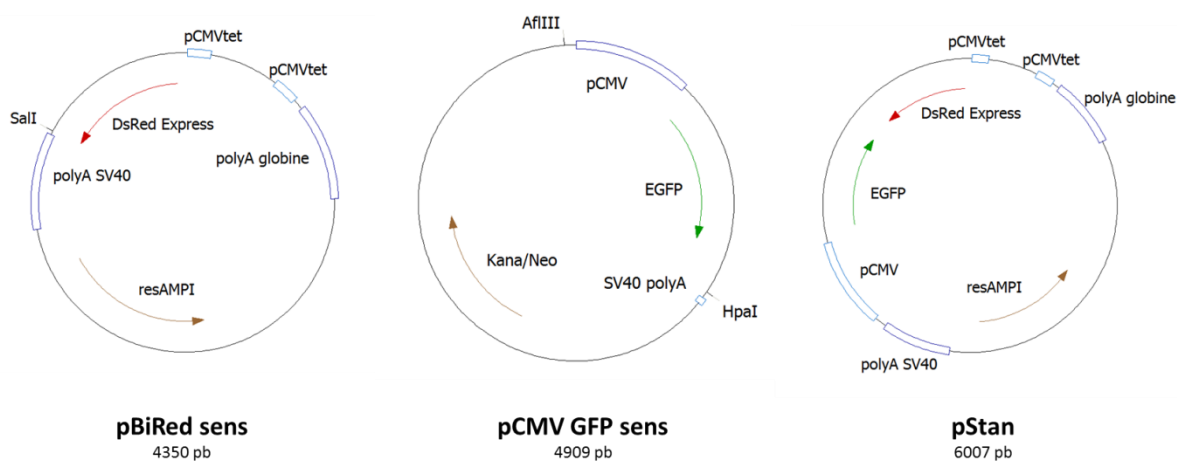


**Figure M5 : Carte des constructions pCMV**

Carte de 3 constructions. pNirFP-N et les deux constructions dérivées pCMV GFP sens et antisens

## 1.3 – La construction pStan

Pour la recherche d'un silencing en *cis* par un ARN antisens, la construction pStan a été réalisée. La construction pStan est issue du plasmide pBiRed sens linéarisé entre la séquence codant la DsRed et la séquence de polyadénylation dans laquelle a été inséré un fragment de la construction pCMV GFP sens comportant la séquence codante de la GFP et le promoteur pCMV (Fig. M6). L'insertion de ce fragment est fait en orientation antisens avant de réaliser une construction pCMV GFP sens DsRed antisens pBi qui, en présence de doxycycline dans milieu de culture, permet l'expression d'un ARN antisens en *cis*.



**Figure M6 : Carte de la construction pStan et des constructions dont il y est issu**

Sur les cartes de pBiRed sens et pCMV GFP sens, les sites des enzymes de restriction utilisés pour la linéarisation de pBiRed sens (SalI) et l'excision du fragment pCMV GFP sens (AflIII et HpaI) sont mentionnés.



## 1.4 – Le Binding Element PRC2

Pour tester si le recrutement d'un complexe modificateur de la chromatine pouvait augmenter l'effet de l'action de l'antisens, un élément permettant ce recrutement a été ajouté à plusieurs constructions comme pBiRed sens GFP antisens ou pStan. La séquence utilisée est basée sur les travaux de l'équipe de Shankar Balasubramanian qui a mis en évidence la séquence minimale permettant le recrutement du complexe PRC2 de l'ARN non codant HOTAIR (Wu *et al*, 2013). Cet élément a été introduit sous forme d'oligonucléotides (Eurogentec) en la flanquant de sites de restriction (Fig. M7).

Oligo Sens

ACC CTTAAG CCTTGCCTAAATCTTCGGACGTCATCCCCCTACCCCTCACCTCTCTCCCTCGGGTCTCAATGTCTGCCGCCGCTCTCTCTCCCCGC TCCTGGA CTTAAG GTA

Oligo Antisens

TAC CTTAAG TCCAGGA GCGGGGAGGAAGGAGAGCGGCGGCAGACATTGAGACCCGAGGGAGAGAGGTGAGGGGTGAGGGGATGACGTCCGAAGATTAGGCAAGG CTTAAG GGT

AflIII PfoI PRC2 Binding Element

### Figure M7 : Séquence de l'élément de recrutement du complexe PRC2

La séquence de l'élément issue de l'article de l'équipe de S. Balasubramanian apparaît en rouge. Les séquences en vert correspondent aux séquences permettant la création du site pour l'enzyme de restriction AflIII et les séquences en bleu correspondent aux séquences permettant la création du site pour l'enzyme de restriction PfoI.

## 2 - Les siARNs

Dans notre travail, plusieurs siARN ont été utilisés pour inhiber l'expression de cibles d'intérêt. Les siARN sont synthétisés chimiquement par MWG Biotech ou Eurogentec et fournis lyophilisés sous forme double brin. Les siARN sont mis en solution à une concentration finale de 100  $\mu$ M (solution stock) dans une solution contenant 100 mM de sel comme le siMAX 1X (MWG / Operon). Pour les transfections on utilise comme source une solution à 10  $\mu$ M ou moins si nécessaire, afin de pouvoir mesurer la quantité nécessaire de manière précise ; typiquement l'ajout de 1  $\mu$ L de solution siARN à 10  $\mu$ M dans le mélange de transfection d'un puits d'une plaque de 12 puits (2 mL de milieu de culture) donne une concentration finale de 5 nM. La notion de concentration est utilisée ici par simplicité pour définir la quantité de siARN utilisée mais n'a pas de signification réelle puisque dans la



plupart des techniques de transfection, les acides nucléiques sont sous forme d'agrégats et non d'une solution homogène et que l'on ignore l'efficacité de transfection des cellules.

**Tableau 1 : Description des siARNs utilisés**

Nom du siARN	Espèce	Source	Nature de la cible	Séquence (5' → 3')
siCD9	Souris	équipe	ARNm	UUCUCUUGCUCGAAGAUGC
siAGO2A	Homme	équipe	ARNm	CGAAGUGCAAGUUUCAACA
siDICER	Homme	équipe	ARNm	UGCUUGAAGCAGCUCUGGAdTdT
siPKR	Homme	Nejepinska <i>et al</i> , 2012	ARNm	GCGAGAAACUAGACAAAGU
siRNase L	Homme	Broad Institute, shRNA TRCN0000355951	ARNm	AUAGUGAUGUGGAGGCUAUUA
siDsRedB	Homme	équipe	ARNm	UGUAGAUGGACUUGAACUC
siPAF1	Homme	Nagaike <i>et al</i> , 2011	ARNm	AAGCAGCAGTTTACCGAGGAA
siCtr9	Homme	Nagaike <i>et al</i> , 2011	ARNm	GCACGUUAUGAUGGCAAUU

## 3 - Extraction des ARNs et RT-qPCR

### 3.1- Extraction des ARNs

Pour l'extraction des ARNs analysés en RT-qPCR, nous utilisons et suivons le protocole du RNeasy Mini Kit de Qiagen. Le protocole est réalisé en utilisant des boîtes de 60 mmensemencées à  $4.10^5$  cellules 72h plus tôt en présence ou absence de doxycycline.



## 3.2 – Reverse transcriptase

Pour l'analyse des ARNs en qPCR, il est nécessaire que les ADN complémentaires de ces derniers soit obtenus par reverse transcriptase. Pour réaliser cela, il est utilisé le High Capacity cDNA Reverse Transcription Kit de Applied Biosystems. Pour chaque échantillon d'ARN extrait réverse transcrit, un témoin négatif sans reverse transcription est effectué pour les analyses en qPCR. De façon standard, la reverse transcriptase est réalisée avec les Random Primers fournis avec le kit.

Néanmoins, dans le cas de l'étude de pStan une reverse transcription spécifique a été faite pour différencier dans l'analyse les transcrits GFP sens issus du promoteur CMV et antisens issus du promoteur pBi.

Séquences des primers utilisés pour reverse transcription spécifique :

Primer RT transcrit pCMV → 5' – GCTACCGGACTCAGATCCAC – 3'

Primer RT transcrit pBi → 5' – ACTCAGATCTCGAGCTCAAG – 3'

## 3.3 – Analyse par qPCR

Les PCR en temps réel ont été effectuées dans un appareil Light Cycler (Roche) dans un volume final de 10 µL en utilisant le SYBR green comme agent intercalant. Les réactifs utilisés sont ceux du « kit Maxima® SYBR Green qPCR Master Mix (2X) » (Fermentas). Le « master mix » contient la Maxima® Hot Start Taq DNA polymérase, le SYBR® Green I et les dNTPs dans un tampon PCR optimisé. Selon le protocole de Fermentas, on distribue 8 µL/capillaire du mélange commun sont distribués auxquels sont ajoutés 2 µL de l'ADNc (produit de transcription inverse non dilué) dans le capillaire. Les capillaires sont centrifugés à 800 tours/min pendant 20 secondes.

Le protocole standard de qPCR est le suivant :

- Pré-incubation 95° → 10 min
- 40 cycles d'amplification
  - 95° → 15 secondes
  - 51° → 30 secondes
  - 72° → 30 secondes
- Courbe de fusion après le dernier cycle



## Matériels et méthodes

---

L'analyse des données est faite en utilisant la dérivée seconde de la fluorescence pour définir le cycle auquel émerge le signal (cycle du seuil « Ct »).

**Tableau 2 : Primer utilisé pour RT-qPCR**

Nom du primer	Séquence (5' → 3')
Commun pCMV	CTCGAGCTCAAGCTTCGAATTCGT
Commun pBi	GGCCGCTCTAGAACTAGTGG
Commun pBiRedGFP as	CTCGTTTGGGATGGGTACAG
GFP sens	ACCTTGTGGCCGTTTACGTC

Un seul primer GFP sens est utilisé car la détection de l'antisens dans les clones HEK pBiRed sens GFP antisens est possible avec le couple de primer Commun pBiRedGFPas et primer GFP sens du fait que le primer Commun pBiRedGFPas se situe en antisens entre la partie codante GFP et la séquence de polyadénylation. Dans le cas de pStan, l'antisens est détecté grâce à la reverse transcription spécifique.



## Résultats

Afin de comprendre et décomposer au mieux les mécanismes d'action des ARN antisens, la stratégie adoptée est d'utiliser un modèle totalement artificiel et simple que l'on pourra complexifier par la suite. Ce modèle est composé de constructions indicatrices permettant l'expression de protéines fluorescentes. Le contexte « antisens » est réalisé par l'inversion de l'orientation du cadre de lecture permettant ainsi la formation d'un ARN antisens complémentaire de l'ARN messager de la protéine fluorescente en question. Pour pouvoir suivre l'expression de l'antisens, cet ARN antisens complémentaire a aussi été introduit en aval de la région codante d'une autre protéine fluorescente. Ces constructions indépendantes permettent d'explorer les possibilités d'une régulation en *trans*, tandis que le contexte en *cis* nécessite des constructions spécifiques avec une transcription convergente.

Une première partie du travail a été réalisée en transfection transitoire afin de s'affranchir du contexte génomique et au moins en partie de l'organisation de la chromatine tandis que dans un deuxième temps la régulation de l'expression de ces constructions a été étudiée après intégration dans le génome.

### 1 – Effet en transitoire

#### 1.1- Première observation de l'effet d'un antisens

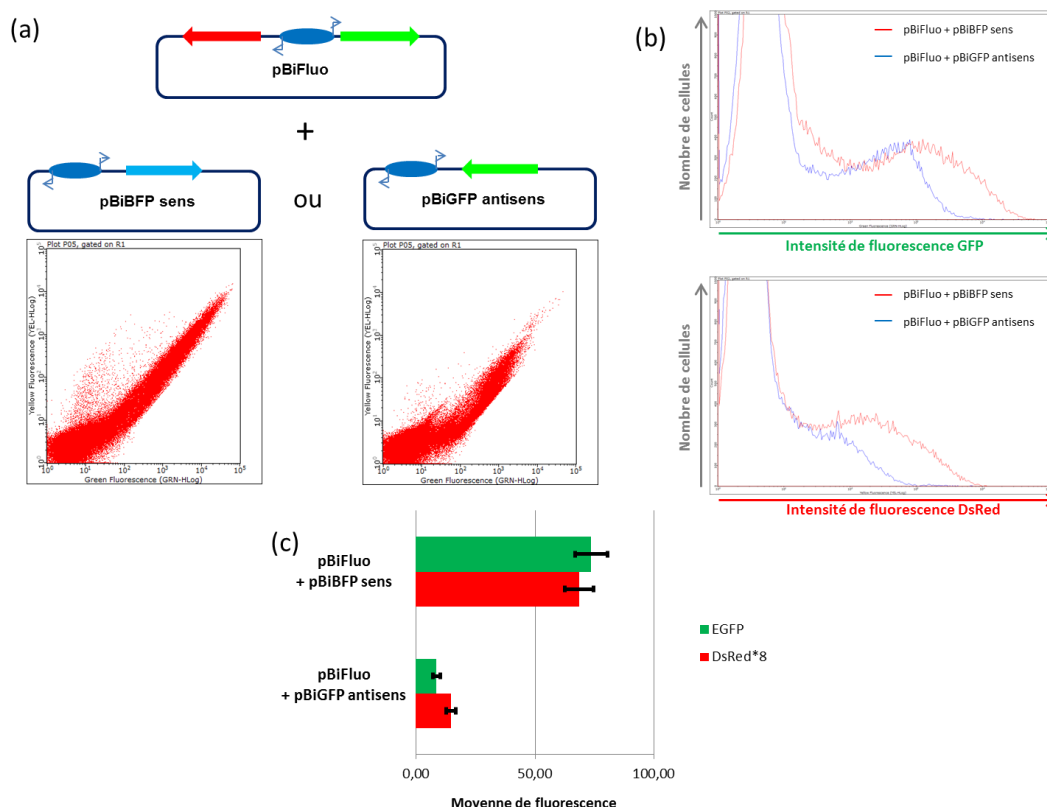
Mes premières expériences ont utilisé comme « cible » la construction indicatrice pBiFluo qui exprime les deux protéines fluorescentes GFP et DsRed sous le contrôle du promoteur bidirectionnel régulé par la tétracycline. Cette construction a été co-transfectée avec un vecteur d'expression pour l'ARN GFP antisens (Fig. 17). Sur les valeurs moyennes de fluorescence on observe une diminution de plus 80% de la fluorescence GFP par rapport au contrôle (un vecteur d'expression d'organisation similaire et permettant d'exprimer la protéine fluorescente BFP – qui n'est pas excitable dans les cytomètres que j'ai utilisé). Grâce à l'analyse par cytométrie, on peut observer que l'ensemble de la population des cellules exprimant la GFP ne réagit pas de la même façon à la présence de l'antisens (Fig. 17b). La réponse à l'antisens va être bimodale et semble ne pas affecter de la même façon





# Résultats

les forts et les faibles expresseurs de la GFP. Ainsi, alors que les faibles expresseurs se déplacent au niveau de l'autofluorescence, les plus forts expresseurs eux se déplacent au niveau des moyens expresseurs ce qui conduit à un épaulement. Cela suggère que l'efficacité du silencing (et peut-être le mécanisme sous-jacent) diffère en fonction du niveau d'expression.



**Figure 17 : Effet d'une construction antisens sur une construction sens en transfection transitoire**

Expériences réalisées par transfection au phosphate de calcium des cellules HEK 293 3G Tet-on (0,25 µg de chaque plasmide pour un puits d'une plaque 12 puits) et analyse 48h après transfection en cytométrie en flux (GUAVA). (a) Présentation schématique des constructions utilisées. Les résultats de l'analyse par cytométrie en flux sont présentés sous forme de nuage de points où chaque point correspond à la fluorescence d'une cellule. L'axe X représente le canal de détection de la GFP et l'axe Y le canal de détection de la DsRed. (b) Histogrammes de distribution des cellules selon l'intensité de fluorescence. La courbe rouge représente la condition sans la construction « antisens » et la courbe bleue représente la condition avec la construction « antisens ». (c) Histogrammes des valeurs moyennes de fluorescence pour les deux fluorescences GFP et DsRed dans les deux conditions avec et sans la construction « antisens ».

Le fait marquant de cette expérience préliminaire a été que sur cette cible bi-fluorescente, l'expression de la DsRed est tout aussi affectée par la présence de l'antisens que la cible directe GFP. Bien que les gènes de la GFP et de la DsRed soient faiblement apparentés un effet antisens direct du transcrit GFPas sur la DsRed est improbable et peut être exclu par une expérience contrôle (voir plus bas). Comme l'expression des deux protéines



fluorescentes est sous le contrôle du même promoteur, l'observation d'un « co-silencing » de la DsRed pouvait suggérer que le transcrit antisens de la GFP avait un impact au niveau transcriptionnel. Compte tenu du caractère transitoire de l'expérience et de l'absence de séquence spécifique recrutant des acteurs de la régulation transcriptionnel sur le transcrit antisens, une telle régulation au niveau transcriptionnelle était inattendue mais aurait pu refléter une régulation forte liée aux niveaux d'expression très élevés des transcrits sens et antisens dans les cellules les plus faciles à analyser – les forts expresseurs.

## 1.3 – Un effet transcriptionnel de l'antisens ?

Pour tester cette hypothèse, le même type d'expérience a été faite mais cette fois-ci en constructions séparées. En plus de la co-transfection avec une construction sens et antisens une construction contrôle exprimant une autre protéine fluorescente est ajoutée. Celle-ci ne devrait pas être affectée par une régulation transcriptionnelle liée à une action de l'antisens sur sa cible directe. Dans le cas présenté en figure 18, le contrôle est la construction pBiGFP sens avec laquelle on co-transfecte soit pBiNir sens soit pBiNir antisens soit le couple sens / antisens. Comme dans toutes les expériences de co-transfection la première observation est une compétition entre plasmides qui se traduit par une diminution de l'expression ici clairement visible sur pBiNir. Comme ces plasmides ont le même promoteur on peut supposer qu'il s'agit d'une compétition pour le transactivateur tétracycline. Néanmoins, d'autres interactions plus complexes semblent possibles car la co-transfection de trois plasmides n'augmente pas le phénomène. En présence d'une construction antisens on observe une diminution de l'expression de la cible directe (Nir) mais aussi de la construction contrôle (GFP). Ce silencing sur la construction contrôle pBiGFP sens n'est pas dû à un effet de la construction antisens Nir et nécessite le couple sens/antisens. Il est donc probable que la formation du duplex sens/antisens est indispensable pour la mise en place de cette régulation. Il est possible que l'effet observé dépende d'un mécanisme spécifique du promoteur qui dans cette expérience est commun à toutes les constructions. Cependant, le même type de réponse a été observé en présence de constructions avec différents promoteurs. Ceci nous a orienté vers d'autres mécanismes qu'une régulation transcriptionnelle et en particulier des mécanismes non spécifiques.



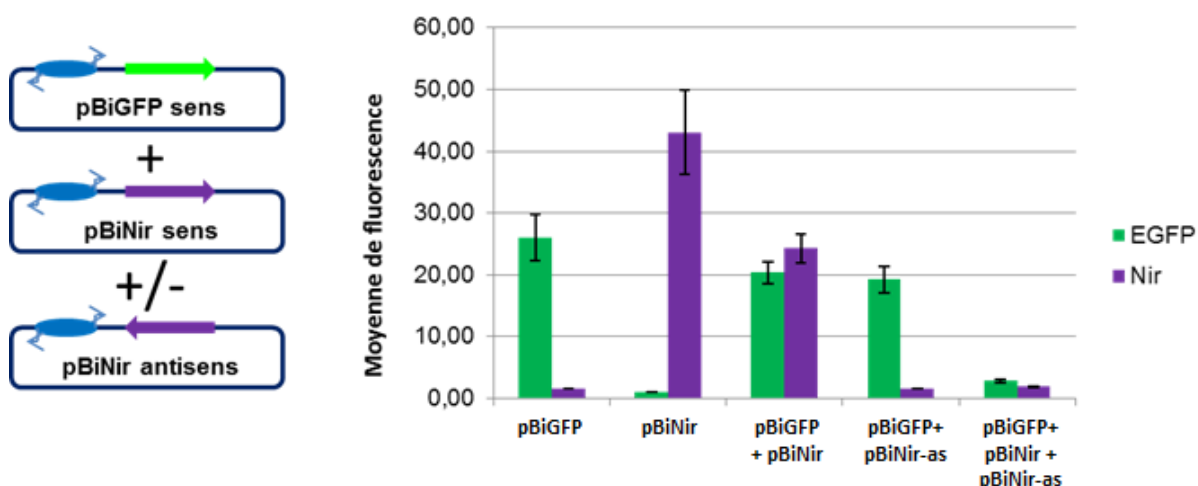


Figure 18 : Effet de la construction antisens sur une construction sens et un contrôle avec construction séparée

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de chaque plasmide et analyse 48h après transfection en cytométrie en flux (FACS Calibur). Présentation de l'expérience et histogramme des valeurs moyennes de fluorescence obtenues après analyse en cytométrie en flux. Transfection de 5 combinaisons différentes de plasmides présentées en abscisse de l'histogramme.

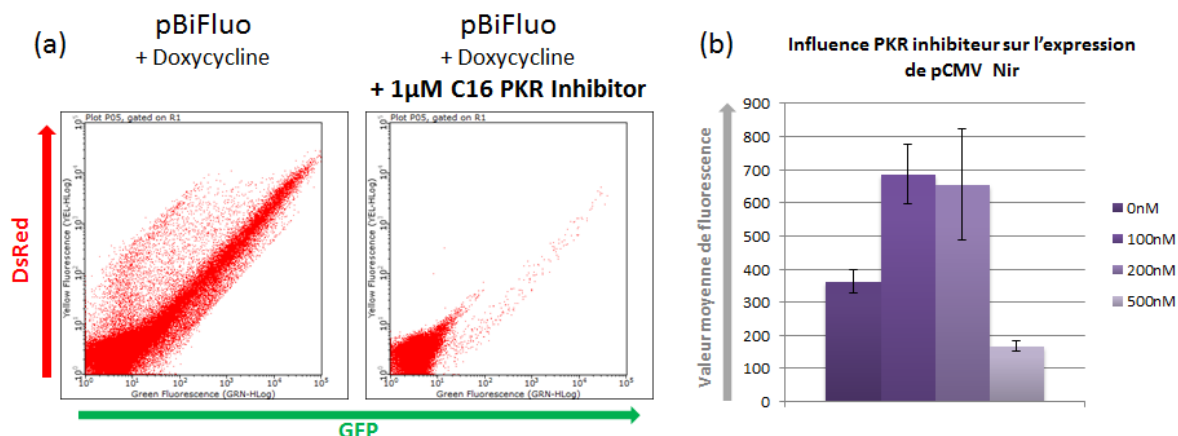
## 1.4 – Etude de la réponse non spécifique PKR-dépendante

L'expression d'ARN double brin, pour autant qu'ils soient présents dans le cytoplasme, est susceptible d'induire une réponse de type interféron. Des travaux menés par l'équipe de Petr Svoboda ont montré récemment une activation partielle de la voie interféron impliquant la protéine kinase PKR en réponse à l'expression d'un ARN double brin de grande taille (Nejepinska *et al*, 2014). Leurs résultats sont similaires à ceux présentés ici et montrent que la présence d'ARN double brin induit le silencing d'une construction indicatrice co-exprimée en transitoire. L'activation de la protéine PKR par la présence d'un duplex d'ARN dans ce modèle transitoire pourrait donc expliquer les réponses non spécifiques observées. Pour inhiber la protéine kinase PKR, deux types d'inhibiteurs ont été utilisés. L'inhibiteur chimique C16 (Jammi *et al*, 2003) et un siARN ciblant PKR (Nejepinska *et al*, 2014).

### 1.4.1 – Inhibition de PKR par l'inhibiteur chimique C16

L'inhibiteur chimique C16 est un dérivé de l'imidazole-oxidole qui a comme propriété de bloquer l'accès au site de fixation de l'ATP nécessaire à l'activité kinase de PKR. Cette

molécule bloque donc l'autophosphorylation de la protéine PKR et en conséquence son effet inhibiteur de la traduction.



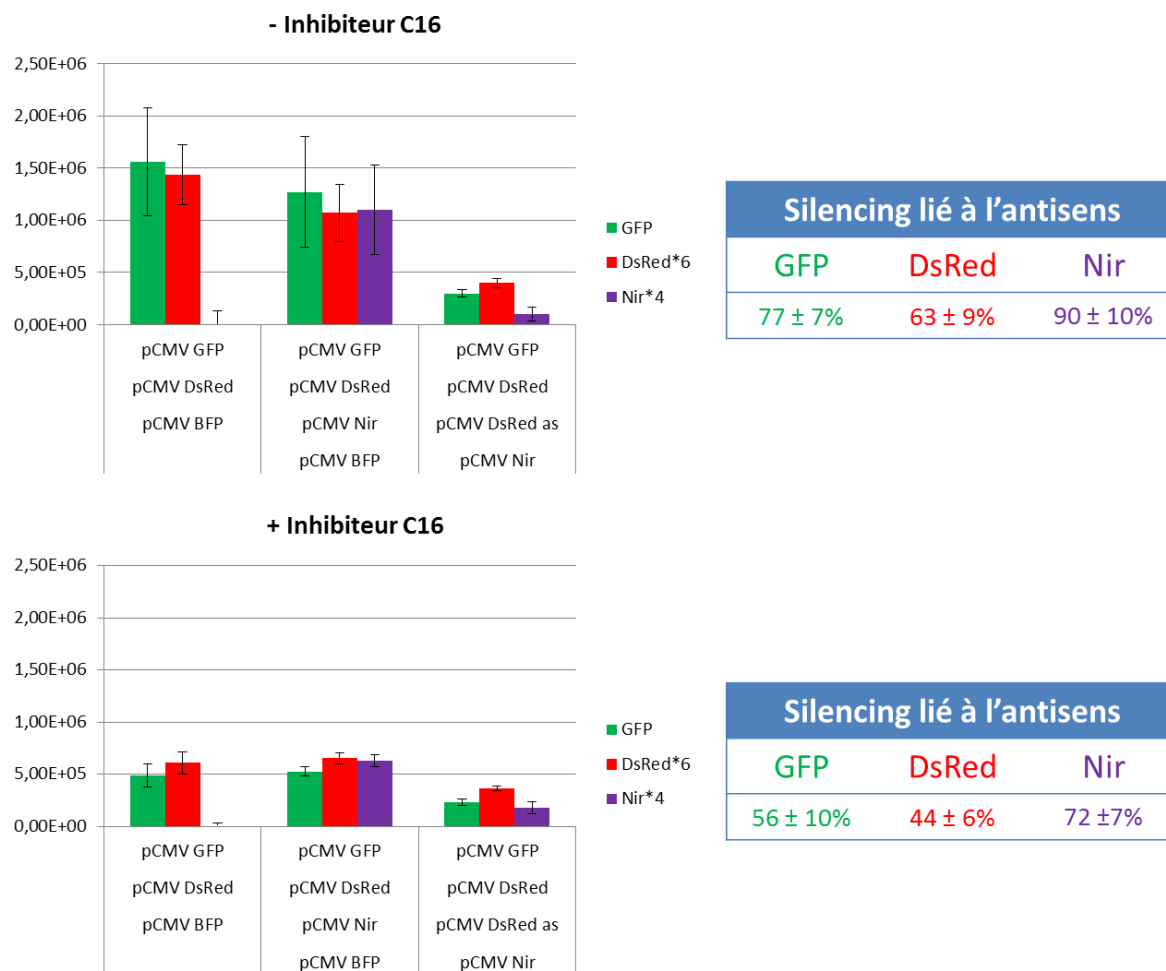
**Figure 19 : Effet de l'inhibiteur chimique de PKR, C16**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de chaque plasmide et analyse 48h après transfection au phosphate de calcium en cytométrie en flux (GUAVA). Ajout de l'inhibiteur C16 le jour de la transfection. (a) Profil de fluorescence des cellules transfectées par le plasmide pBiFluo en présence ou absence de l'inhibiteur. (b) Dose réponse de l'inhibiteur C16 (0, 100 nM, 200 nM et 500 nM), test réalisé sur la transfection du plasmide pCMV Nir seul.

Les premiers tests réalisés avec cet inhibiteur se sont montrés peu concluants. A la concentration conseillée par le fournisseur (1 µM), l'expression des constructions et la viabilité des cellules sont très fortement diminuées (Fig. 19a). Pour déterminer la dose d'inhibiteur adéquat pour réaliser ces tests, une dose réponse d'inhibiteur a été faite sur des simples transfections à des doses croissantes d'inhibiteur allant de 0 à 500 nM (Fig. 19b).

A faible dose (100 nM ou 200 nM), on observe une forte augmentation de l'expression de la construction pCMV Nir transfectée, tandis qu'à 500 nM une diminution de l'expression est associée à une baisse de la viabilité. Ces données suggèrent que PKR pourrait être impliquée dans le niveau de base d'expression des plasmides transfectés. Ceci va dans le sens d'autres travaux réalisés par l'équipe de Petr Svoboda qui montrent que les constructions plasmidiques habituelles génèrent de nombreux transcrits convergents issus de promoteurs cryptiques présents dans la séquence des plasmides. Ceci conduit à une inhibition du niveau basal de l'expression (Nejepinska *et al*, 2012). On peut penser que cette inhibition par des transcrits « illégitimes » participe aussi aux effets de compétitions entre plasmides dans les expériences de co-transfection. L'inhibition de la protéine PKR semble permettre une levée de ces mécanismes et donc une augmentation de l'expression des constructions transfectées. La dose choisie pour le reste des expériences est de 200 nM.





**Figure 20 : Effet de l'inhibiteur chimique de PKR, C16, sur le silencing par un antisens**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de chaque plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). L'inhibiteur de la protéine PKR est ajouté le jour de la transfection et maintenu lors du changement de milieu 24h après transfection. Les histogrammes représentent les valeurs moyennes de fluorescence dans chacune des conditions avec ou sans construction antisens en présence ou non de l'inhibiteur de la protéine PKR. Les tableaux représentent le pourcentage de diminution d'intensité de fluorescence entre les conditions avec et sans la construction antisens. Pour une meilleure visualisation des données, les valeurs de fluorescence de la DsRed et Nir sont multipliées respectivement par 6 et 4.

Afin de déterminer si l'ajout de cette construction « antisens » permet de déclencher un silencing non spécifique PKR-dépendant, l'inhibiteur chimique C16 a été testé dans des conditions de silencing par un antisens. Ces derniers résultats sont comparés aux mêmes conditions sans l'inhibiteur (Fig. 20). En absence d'inhibiteur (Fig. 20), on retrouve les données déjà observées avec la présence d'un fort silencing sur l'expression de l'ensemble des constructions co-transfectées, aussi bien la cible directe de l'antisens (DsRed), que les contrôles (GFP et Nir). En présence de l'inhibiteur, on s'aperçoit que les niveaux d'expression sont plus faibles de plus d'un facteur 2 par rapport aux valeurs en absence d'inhibiteur et



cela pour toutes les fluorescences. Cependant, si l'on observe les valeurs du silencing celle-ci diminuent de 20% pour la cible directe comme pour les cibles indirectes. Il est donc possible que PKR soit impliquée dans le silencing non spécifique induit par l'expression d'un antisens. Cette conclusion est cependant fragile puisque dans ces expériences l'inhibiteur C16 a un fort impact sur la viabilité des cellules et sur l'efficacité de la transfection. Une explication alternative serait que la diminution du silencing soit due à transfection moins efficace de la construction antisens.

La forte toxicité sur les cellules de C16 et son effet sur l'efficacité de la transfection ne permettent donc pas de conclure sur l'implication de PKR dans les réponses observées.

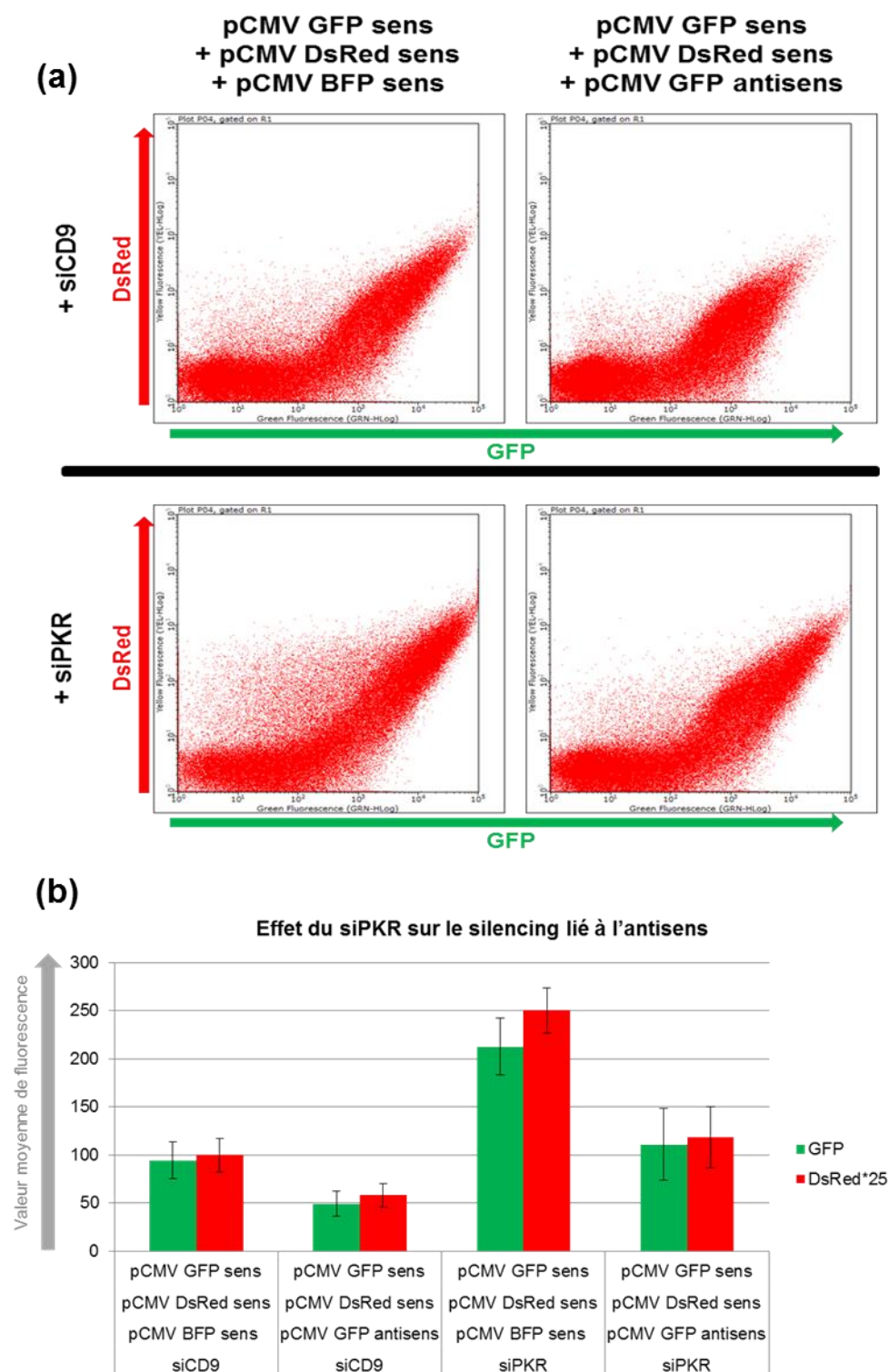
## *1.4.2 – Inhibition de PKR par l'utilisation d'un siARN*

L'utilisation d'un siARN ciblant le transcrite PKR va permettre moins de toxicité cellulaire et de ne pas perturber la transfection. Toutefois, l'inhibition permise par cette approche est potentiellement moins complète car elle dépend de l'efficacité du silencing induit par le siARN et de l'efficacité de transfection. Il est important de noter que des travaux réalisés précédemment dans l'équipe ont montré la possibilité d'induire un silencing par siARN dans pratiquement toutes les cellules, du moins dans le modèle utilisé ici des cellules HEK 293 tet-on transfectées au phosphate de calcium. La séquence choisie pour cibler la PKR a été dérivée de la séquence de l'ARN en tige-boucle utilisée par l'équipe de Petr Svoboda (Nejepinska *et al*, 2014). De la même façon que pour l'inhibiteur chimique C16, le siARN ciblant le transcrite de la protéine PKR est ajouté au mélange de transfection avec ou sans construction antisens (Fig. 21). En l'absence de construction antisens, l'inhibition de l'expression de PKR conduit à une augmentation de l'expression des plasmides transfectés visible sur les nuages de points (Fig. 21a) et quantifiable par les valeurs moyennes de fluorescence (Fig. 21b). Cette augmentation d'environ 2,5 fois est similaire pour la GFP et la DsRed. A l'opposé de cette implication de la PKR dans l'expression de base des plasmides, le silencing induit par la présence d'une construction GFP antisens n'est pas affecté que ce soit sur sa cible directe (GFP, silencing  $\sim 2x$ ) ou sur le contrôle DsRed (silencing  $\sim 2x$ ). Même si l'on prend en compte un silencing potentiellement incomplet de la kinase PKR, cette différence de réponse entre le niveau de base d'expression et le silencing induit par



l'expression d'un antisens suggère que ce dernier ne dépend pas de manière critique de la voie PKR. Ce résultat est différent de celui observé par l'équipe de P. Svoboda mais il convient de noter que les ARN double brin ne sont pas produits de la même manière dans les deux modèles. En effet, dans les expériences de P. Svoboda c'est le transcrit lui-même qui contient une longue répétition inversée de séquence qui génère un ARN double brin intramoléculaire, à l'image de ce qui a été proposé pour la régulation de l'expression de l'interféron gamma par PKR (Cohen-Chalamish *et al*, 2009). Dans les expériences décrites ici les transcrits sens et antisens sont générés indépendamment et potentiellement à partir de sites de transcription différents, ce qui peut limiter la formation d'ARN double brin ou les restreindre au compartiment nucléaire.





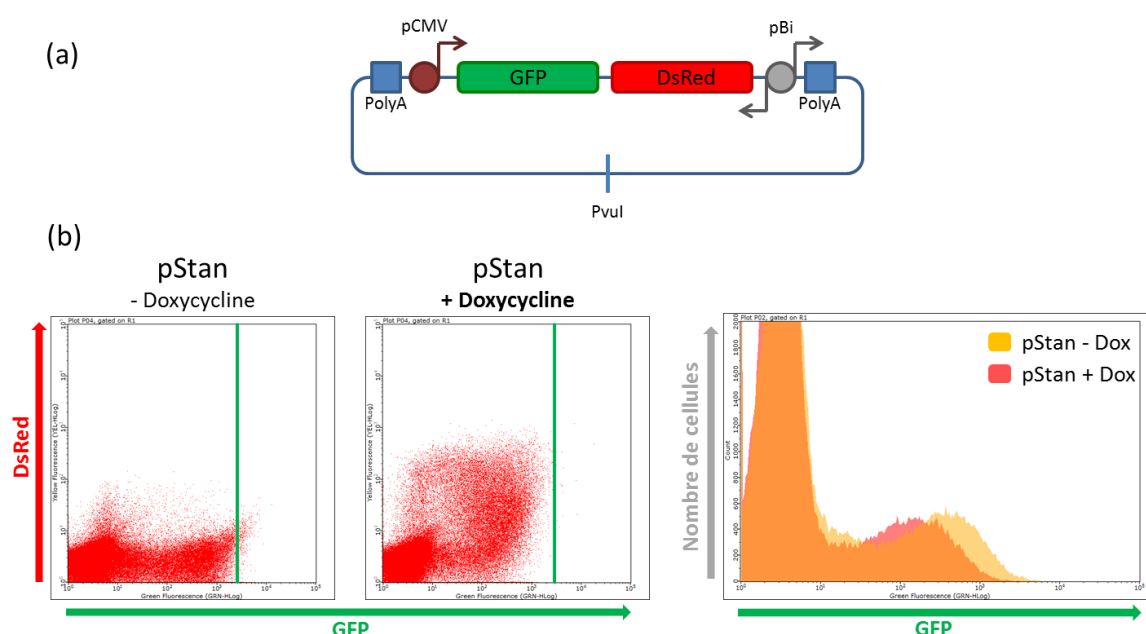
**Figure 21 : Effet du siARN ciblant PKR sur le silencing par un antisens**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de chaque plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). Le siARN est transfecté dans le même mélange que les plasmides à la dose de 5 nM. (a) Profils de diffusion des cellules transfectées par les constructions plasmidiques dans les conditions sans et avec la construction antisens. Test fait avec le siARN contrôle siCD9 (ciblant le transcrite du récepteur CD9 murin) ou avec le siPKR. (b) Histogramme représentant les valeurs moyennes de fluorescence pour la GFP et la DsRed pour chacune des conditions de l'expérience. Pour une meilleure visualisation des données DsRed les valeurs sont multipliées par 25.



## 1.5 – Etude de l'effet d'un ARN antisens exprimé en *cis*

Pour se rapprocher de l'activité des ARN antisens naturels (cisNAT), une construction permettant la formation d'un ARN antisens en *cis* a été créée. Cette dernière est faite à partir d'une construction pCMV GFP sens à laquelle est ajoutée en aval de la GFP la séquence codant la DsRed ainsi que le promoteur inducible à la tétracycline pBi positionnés en antisens (Fig. 22a). Cela donne une construction pCMV GFP sens DsRed antisens pBi qui pour simplifier est appelée pStan. Dans ce contexte, l'ajout de doxycycline va induire l'expression d'un transcrit « DsRed sens GFP antisens » dont le niveau d'accumulation dans le cytoplasme pourra être suivi via l'expression de la DsRed. Afin de maximiser le recouvrement entre les transcrit sens et antisens, les sites de polyadénylation ont été placés au-delà des promoteurs des gènes transcrits en antisens. Le premier test fait sur cette construction a été sa transfection en cellules HEK 293 Tet-on 3G en présence ou absence de doxycycline pour induire l'expression de l'antisens en *cis*. Les données de la figure 22b montrent qu'en présence de doxycycline l'expression de la GFP est affectée avec une diminution de presque dix fois pour les forts expressers (Fig. 22b histogramme). Ce qui montre un effet de l'antisens sur l'expression de la GFP.



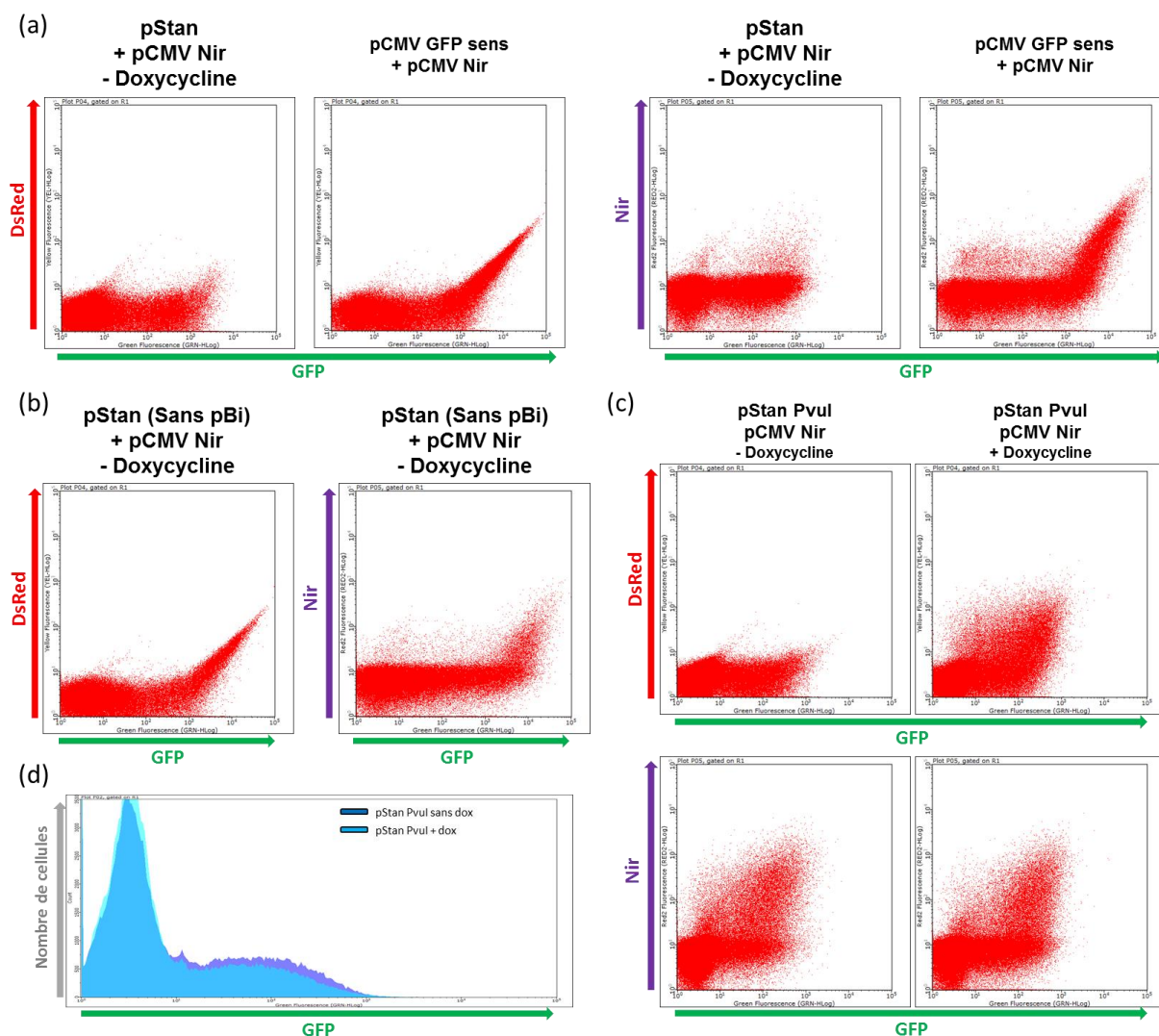
**Figure 22 : Effet de l'induction de l'antisens en *cis* dans la construction pStan**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). L'induction par la doxycycline est faite le jour de la transfection. (a) Schéma de la construction plasmidique pStan. (b) Profils de fluorescence et histogrammes obtenus après analyse en cytométrie en flux des cellules transfectées en présence ou absence de doxycycline. Sur le nuage de points, la barre située à  $3 \times 10^3$  sur l'échelle de fluorescence verte est destinée à faciliter la visualisation du silencing.

Afin de déterminer si ce silencing est spécifique ou non, comme dans le contexte d'un ARN antisens exprimé en *trans*, une construction contrôle a été co-transfectée (Fig. 23). Mais il est apparu que l'expression de la construction pCMV Nir était très faible en présence de la construction pStan et ce même en l'absence de doxycycline (Fig. 23a). Du fait de l'organisation de la construction pStan, un faible taux de base d'expression du promoteur pBi pourrait peut-être suffire à générer de l'ARN double brin compte tenu de la proximité spatiale avec les transcrits sens.

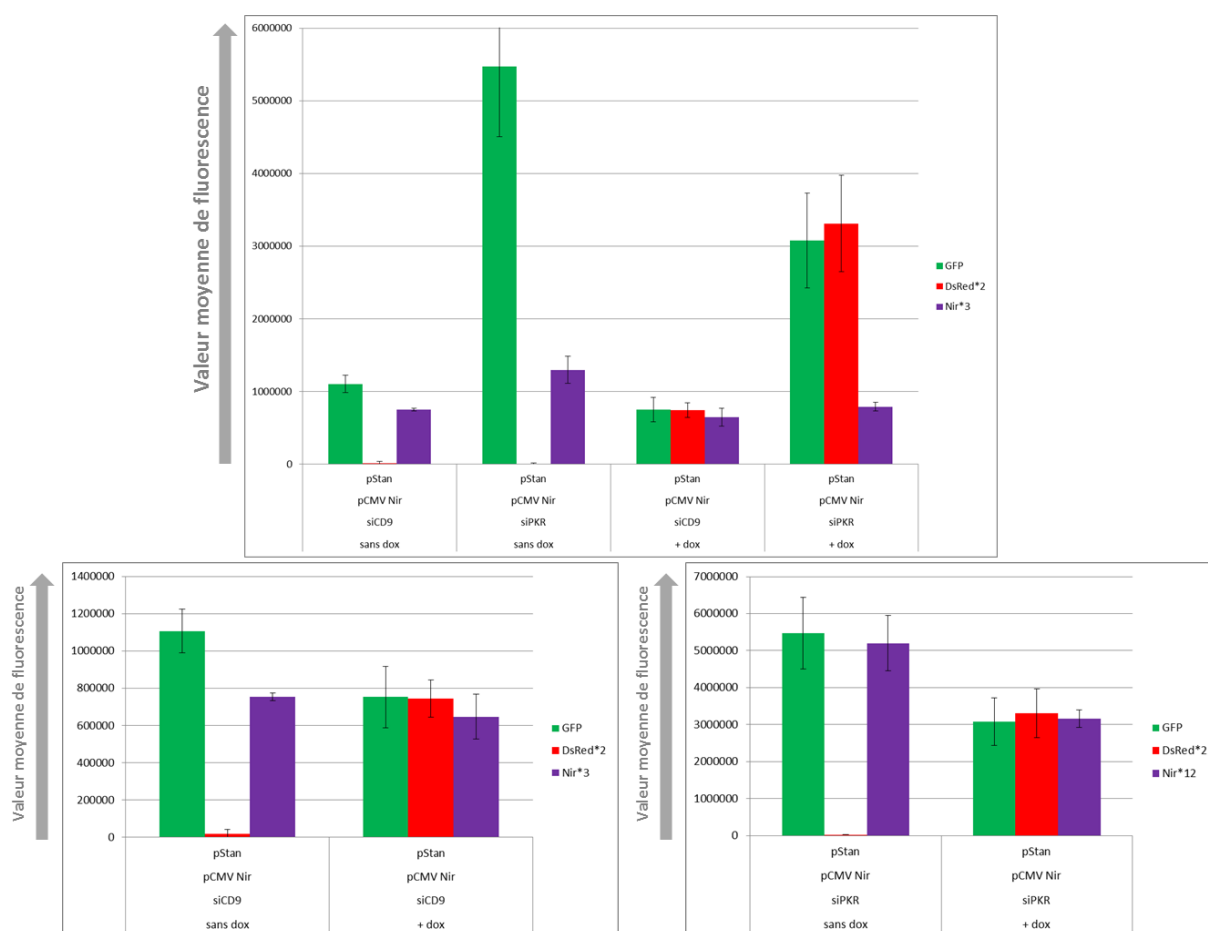
Pour tester cette hypothèse, j'ai délété le promoteur pBi de la construction pStan (Fig. 23b). Cette délétion permet en effet une meilleure expression à la fois en GFP mais également de la construction contrôle co-transfectée pCMV Nir sens. Cela établit une implication des séquences du promoteur pBi dans le silencing observé de la construction CMV Nir. Cependant plusieurs mécanismes sont envisageables, en particulier le promoteur pBi étant bidirectionnel, il peut aussi générer des transcrits de même polarité que l'ARNm de la GFP mais pourrait recouvrir son promoteur. La présence d'un site de polyadénylation en aval de pBi (Fig. 22a) devrait fortement limiter l'expression de tels transcrits mais sans nécessairement les supprimer complètement. Pour tester si de tels transcrits interviennent dans l'expression des constructions transfectées, une linéarisation du plasmide dans le corps du plasmide a été effectuée (PvuI, Fig. 22a). La co-transfection du plasmide linéaire avec la construction contrôle pCMV Nir sens montre que cette dernière est beaucoup moins affectée dans son expression que par le pStan circulaire. Ceci contraste avec l'expression de la GFP qui, elle, reste similaire à celle du pStan circulaire. Ces données laissent à penser que les deux expressions ne sont pas soumises au même type de régulation et que pCMV Nir sens est soumis à un silencing non spécifique lié à des transcrits originaires du corps du plasmide. Le transcrit GFP sens serait, lui, soumis à une régulation plus spécifique par l'antisens GFP. Cependant, l'induction de l'expression de la DsRed et donc de l'antisens GFP n'induit pas de silencing supplémentaire sur la GFP ni d'effet sur l'expression de CMV Nir (Fig. 23c, d) ce qui pourrait remettre en cause cette interprétation.





**Figure 23 : Effet de l'expression de la construction pStan sur celle d'un plasmide contrôle (pCMV Nir)**  
 Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). L'induction par la doxycycline est faite le jour de la transfection. (a) Profil de fluorescence des cellules transfectées par pStan ou pCMV GFP +/- pCMV Nir. (b) Profil de fluorescence de pStan après la délétion du promoteur pBi. Il est co-transfecté avec pCMV Nir comme contrôle. (c) Profil de fluorescence de pStan linéarisé par Pvul. Le plasmide linéarisé est co-transfecté avec pCMV Nir en absence ou présence de doxycycline. (d) Histogramme de la fluorescence de la GFP pour les cellules transfectées par la construction pStan linéarisée par Pvul en absence et présence de doxycycline.

Afin de clarifier les mécanismes impliqués dans ces différents silencings, j'ai testé l'implication de PKR dans l'expression de pStan. Comme discuté plus haut, l'utilisation des inhibiteurs chimiques de PKR comme C16 est inadaptée à ces expériences de transfection et donc j'ai utilisé une co-transfection avec le siARN ciblant PKR (Fig. 24).



**Figure 24 : Effet de l'inhibition de PKR sur le silencing induit par la construction pStan**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). Le siARN est transfecté dans le même mélange que les plasmides à la dose de 5 nM. L'induction par la doxycycline est faite le jour de la transfection. Histogrammes des valeurs moyennes de fluorescence des cellules après analyse de cellules transfectées en cytométrie en flux. L'histogramme principal présente les données avec ou sans le siPKR sur la même échelle. Les deux autres histogrammes présentent les mêmes données mais cette fois-ci sous leurs échelles respectives. Pour une meilleure visualisation des données, les valeurs de moyennes de fluorescence DsRed sont multipliées par 2 et celles de Nir par 3 sauf dans le dernier histogramme par 12.

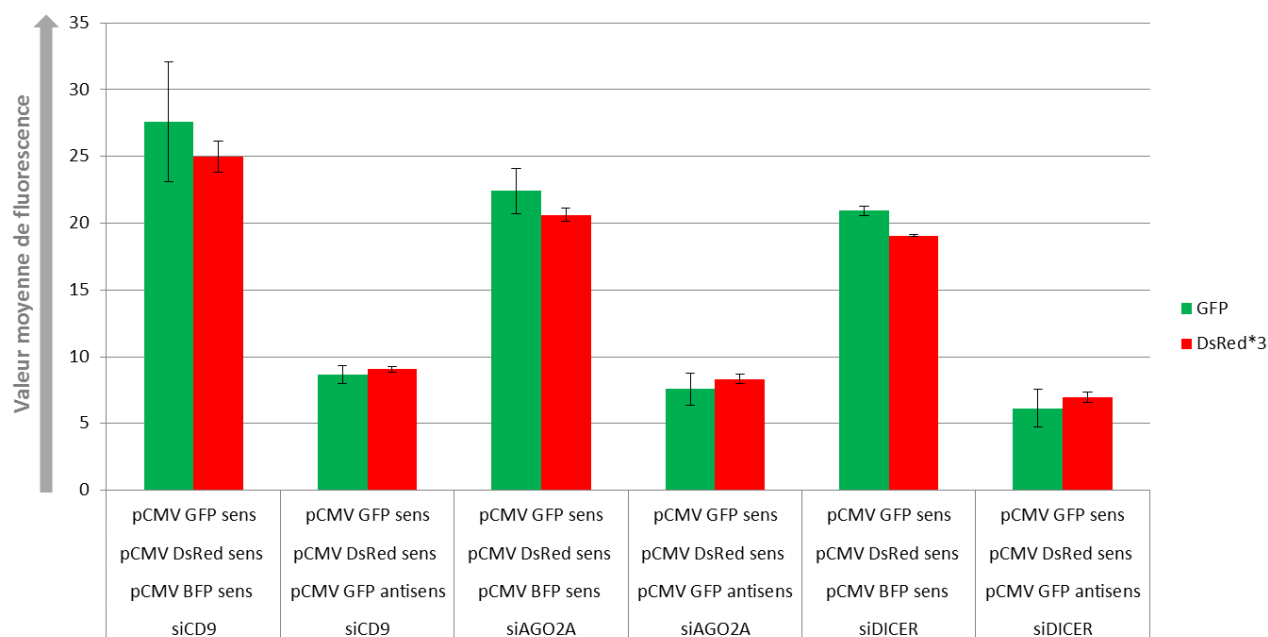
Comme pour les plasmides déjà testés, l'inhibition de PKR va induire une augmentation de l'expression des constructions transfectées. Ainsi en présence du siPKR, l'expression de pStan va être induite d'un facteur 5 (Fig. 24). Cet effet est notablement plus marqué que pour la construction pCMVGFP ce qui pourrait refléter la production d'ARN double brin même en l'absence d'activation de pBi. Néanmoins, et comme c'était le cas pour les constructions dans un contexte en *trans*, le silencing observé en présence de l'antisens n'est pas PKR dépendant (Fig. 24). Ainsi, que l'on soit dans un contexte en *cis* ou en *trans*, le silencing que permet l'expression d'un long ARN antisens qu'il soit spécifique ou non ne semble pas faire appel à un mécanisme PKR-dépendant comme décrit par l'équipe de Petr Svoboda. D'autres mécanismes et partenaires doivent intervenir dans ce phénomène.

## 1.6 – Implication de l'ARN interférence ?

Puisque ces constructions sont susceptibles de générer des ARN double brin, il est possible que ceux-ci interagissent avec Dicer et ainsi induisent une interférence par l'ARN. Il est vrai que cela ne permet pas dans un premier temps d'expliquer la réponse non spécifique observée. Néanmoins, notre modèle pourrait induire des réponses qui ciblent aussi le plasmide de référence, par exemple si des petits ARN ciblent le promoteur CMV ou les régions communes aux différents plasmides. Pour tester cette hypothèse, j'ai utilisé un siARN contre l'ARNm d' Ago2. S'il existe chez les mammifères 4 protéines argonautes, Ago2 est la seule à posséder une activité endonucléolytique et elle est donc l'acteur majeur du silencing sur une cible parfaite. Par ailleurs, j'ai aussi utilisé un siARN ciblant l'ARNm de Dicer ce qui devrait bloquer la maturation des ARN double brin en petits ARN indépendamment de leur prise en charge ultérieure (Fig. 25). L'efficacité du silencing induit par ces deux siARN a précédemment été testée dans l'équipe et est supérieure à 75 % avec un plateau atteint pour une dose de 1 nM (la quantité transfectée est ici de 5 nM comme pour toutes les expériences en présence d'un siARN).

Comme le montre la figure 25, l'inhibition de la voie de l'interférence ARN au niveau de Ago2 ou de Dicer ne modifie que légèrement le niveau d'expression des constructions transfectées à la différence du siPKR (diminution d'environ 20 %, à la limite de la significativité). De plus, aucune modification du silencing associé à l'expression de l'ARN GFP antisens n'est observée que ce soit sur la GFP ou la DsRed. Il est vrai que notre méthode pour inhiber la voie de l'ARN interférence par des siARNs n'est pas efficace à 100% mais néanmoins celle-ci aurait permis si cette voie était impliquée de constater une diminution du silencing. Or il est clair que ce dernier n'est pas affecté fortement ce qui exclut une implication majeure de la voie de l'ARN interférence dans les phénomènes observés.





**Figure 25 : Effet de l'inhibition de Ago2 et de Dicer sur le silencing par un antisens en transitoire**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). Le siARN est transfecté dans le même mélange que les plasmides à la dose de 5 nM. Histogrammes des valeurs moyennes de fluorescence des cellules obtenues après analyse de cellules transfectées en cytométrie en flux. Pour une meilleure visualisation des données, les valeurs de moyennes de fluorescence DsRed sont multipliées par 3.

L'ensemble des données obtenues montre, que l'expression d'un long ARN antisens en transitoire ne donne pas les résultats attendus, que ce soit dans un contexte en *trans* ou en *cis*. Certes on observe un silencing en réponse à l'expression de l'antisens, mais celui-ci est au moins en partie non spécifique et affecte dans des proportions similaires l'expression d'une construction contrôle. Il est envisageable qu'un silencing spécifique de séquence et lié à l'antisens soit mis en place mais son action est donc masquée par l'importance des réponses non spécifiques. Si la production d'ARN double brin est susceptible d'activer une réponse de type interféron avec une activation de la kinase PKR et de la RNaseL, nos données ne montrent pas d'implication forte de cette voie dans le silencing associé à l'expression de l'antisens ; elles confirment par contre un rôle de PKR dans une diminution de l'expression des plasmides transfectés. Ce phénomène décrit par le groupe de P. Svoboda est attribué à la production de transcrits « illégitimes » en particulier dans les séquences de l'ossature plasmidique comme le gène de résistance aux antibiotiques. En générant des ARN double brin ces transcrits activeraient de manière globale PKR et ainsi réduiraient l'expression de tous les plasmides transfectés. Si un tel mécanisme est plausible, il faut



cependant remarquer que les données disponibles sur l'action de PKR favorisent plutôt un modèle d'activation locale par la présence d'une région double brin sur le transcrit à traduire. Par ailleurs, plusieurs données montrent qu'il existe une différence de sensibilité entre les transcrits issus du génome et ceux issus d'une construction épisomale pour ce type de réponse aspécifique impliquant PKR (Terenzi *et al*, 1999 ; Nejepinska *et al*, 2014), ce qui indique que tous les transcrits présents dans la cellule ne sont pas affectés de la même manière. Afin d'explorer plus avant un silencing spécifique induit par l'expression d'un ARN antisens, j'ai donc poursuivi mes travaux en dérivant des clones stables exprimant les constructions présentées ici en transfection transitoire, dans l'idée que les réponses non spécifiques seraient plus faibles. En ce qui concerne le mécanisme des effets non spécifiques, les études présentées ici ne permettent pas de conclure mais néanmoins excluent que PKR et l'interférence ARN soient des acteurs majeurs.

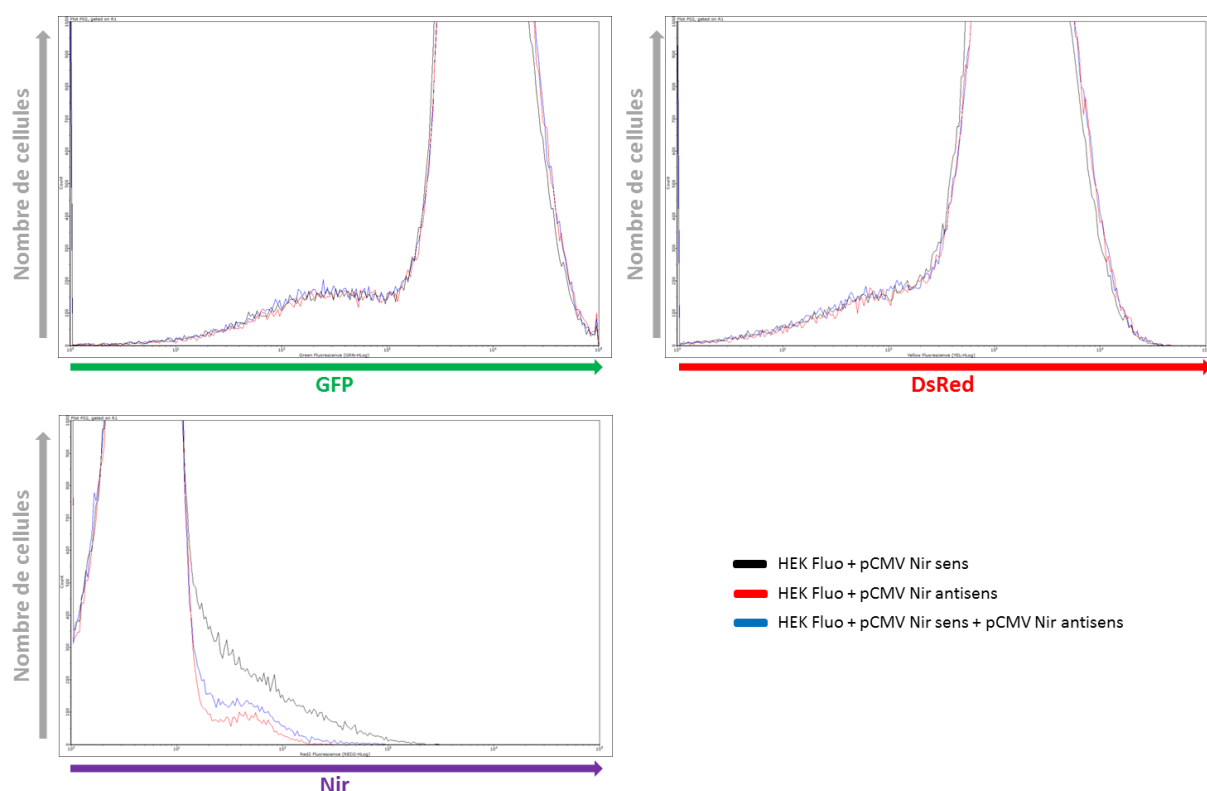
## 2 – Effet en clones stables

L'intégration de nos constructions au sein du génome des cellules HEK 293 Tet-on change de nombreux aspects du modèle. Certes, après sélection d'un clone, le niveau d'expression devient plus homogène, l'intégralité de la population cellulaire exprimant la construction contrairement au modèle transitoire où en général de l'ordre de 10% des cellules expriment significativement la construction. Néanmoins, l'utilisation de clones stables introduit un niveau de complexité supplémentaire puisque l'intégration n'étant pas dirigée, on ignore le locus d'insertion de la construction et si son expression se trouve modifiée ou régulée par des éléments appartenant à ce locus. Une différence anticipée depuis le début du projet est que l'intégration dans le génome doit conduire à l'acquisition d'une conformation chromatinienne complète à la différence de la situation rencontrée dans le cadre d'une expression transitoire à partir d'épisomes. Un deuxième aspect dont l'importance est apparue au cours des essais en transitoire est de diminuer les effets non spécifiques. Ceci se base sur l'hypothèse que le silencing non spécifique déclenché par l'expression des constructions sens et antisens dans la même cellule se comporte de la même façon que le mécanisme PKR-dépendant et d'autres réponses antivirales du même type (Terenzi *et al*, 1999 ; Nejepinska *et al*, 2014). Avant de commencer de plus amples analyses en clone stable



# Résultats

j'ai donc cherché à valider cette hypothèse. Pour cela, des tests ont été faits sur la lignée stable HEK Fluo qui est une lignée ayant intégré la construction pBiFluo (GFP et DsRed sous la dépendance du promoteur inducible à la tétracycline). Cette lignée a été co-transfectée avec un couple de construction sens/antisens pour permettre le déclenchement de la réponse aspécifique (Fig. 26). Comme on peut l'observer, l'ajout du couple de construction CMV Nir sens/antisens en transitoire n'a aucun effet sur l'expression de la construction intégrée. Cela tend à valider notre hypothèse et pose la question de la nature de cette réponse non spécifique. En particulier, la capacité de ce mécanisme à distinguer entre les transcrits issus d'une construction transitoire et génomique est inattendue et implique qu'ils diffèrent par certaines de leurs caractéristiques (couverture protéique, localisation ...).



**Figure 26 : Effet de la réponse non spécifique sur une construction intégrée au génome**

Expériences réalisées dans les cellules HEK Fluo obtenue après intégration de la construction pBiFluo au sein du génome de cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). Trois conditions sont ici testées, la transfection de la construction pCMV Nir sens (courbe noire), la transfection de la construction pCMV Nir antisens (courbe rouge), la co-transfection des deux constructions pCMV Nir sens et antisens (courbe bleue). Les données présentées sont les histogrammes de fluorescence dans les trois canaux (GFP, DsRed ou Nir).

L'élimination de la majorité des effets non spécifiques rend possible l'analyse de l'effet de l'expression d'un long ARN antisens. Deux cas de figure sont donc envisagés pour l'analyse en clone stable au même titre qu'il a été fait pour l'analyse en transfection transitoire. L'analyse de l'effet d'un long ARN antisens exprimé en *trans* en constructions séparées et



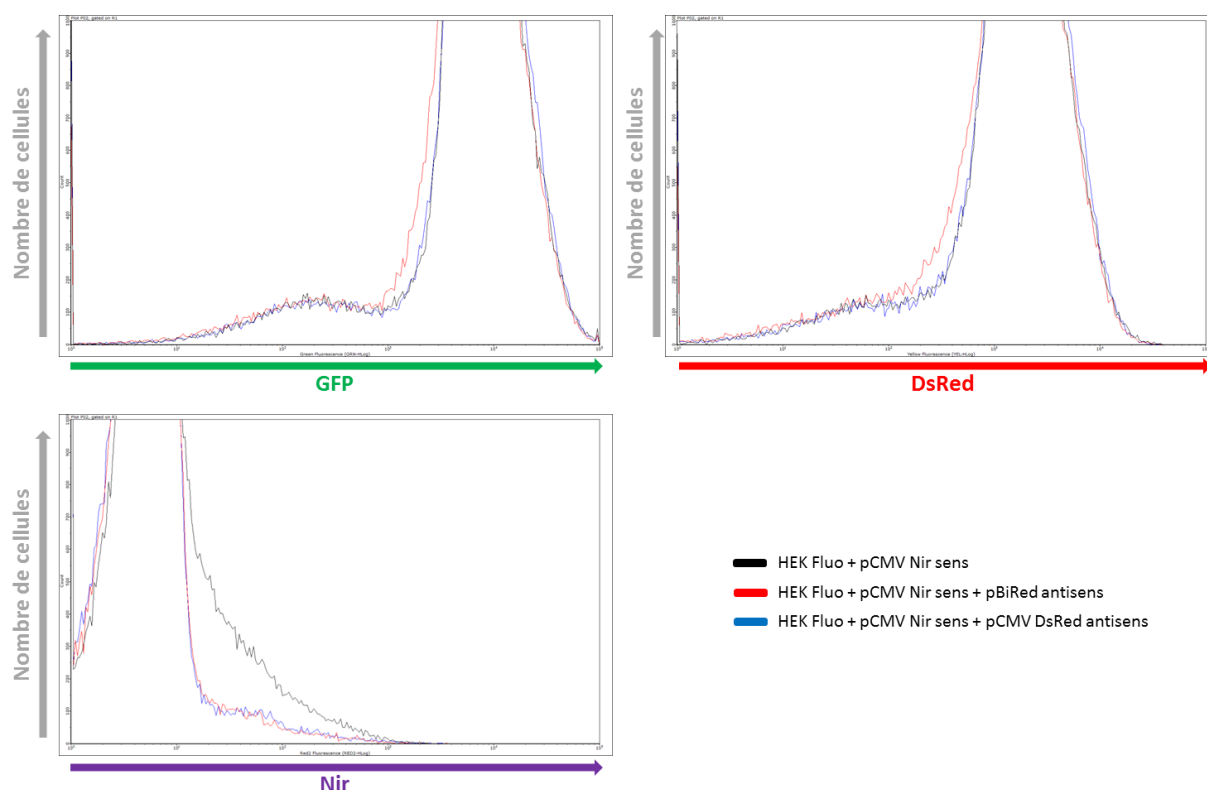
l'analyse en *cis* où l'antisens s'exprime à partir de la même construction que la cible.

## 2.1 –Silencing en *trans* d'une construction intégrée

Ayant choisi dans un premier temps d'exprimer l'antisens en *trans*, il reste une question quant à son expression qui peut être faite à partir d'une construction intégrée ou en transitoire, et si cela a une importance sur l'activité de l'antisens.

### 2.1.1 – L'effet d'un antisens exprimé en transitoire

L'effet d'un ARN antisens exprimé à partir d'une construction transitoire a été la première condition à être testée. Là encore plusieurs choix sont possibles avec l'utilisation de la construction sous dépendance d'un promoteur constitutif ou sous la dépendance d'un promoteur inductible à la tétracycline. Les deux types de constructions ont été testés pour leur effet sur une construction intégrée, ici la lignée HEK Fluo (Fig. 27)



**Figure 27 : Effet d'une construction antisens exprimée en transitoire sur une cible intégrée**

Expériences réalisées dans les cellules HEK Fluo obtenues après intégration de la construction pBiFluo au sein du génome de cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). Les tests sont réalisés soit en absence d'antisens (Courbe noire) ou avec soit la construction pBiRed antisens (Courbe rouge) soit pCMV DsRed antisens (Courbe bleue). Les données présentées sont des histogrammes de fluorescence (GFP, DsRed ou Nir).

Comme on peut le remarquer Figure 27, l'ajout de la construction pCMV DsRed antisens n'a aucun effet sur l'expression de la construction intégrée pBiFluo (GFP/DsRed) et donne le même profil que la transfection en absence d'antisens. En ce qui concerne l'autre construction antisens, pBiRed antisens, qui est sous dépendance du promoteur inductible à la tétracycline, on observe un faible effet de cette dernière sur les deux fluorescences. Mais néanmoins cet effet de la construction antisens reste faible et à la limite de la significativité (silencing < 20%).

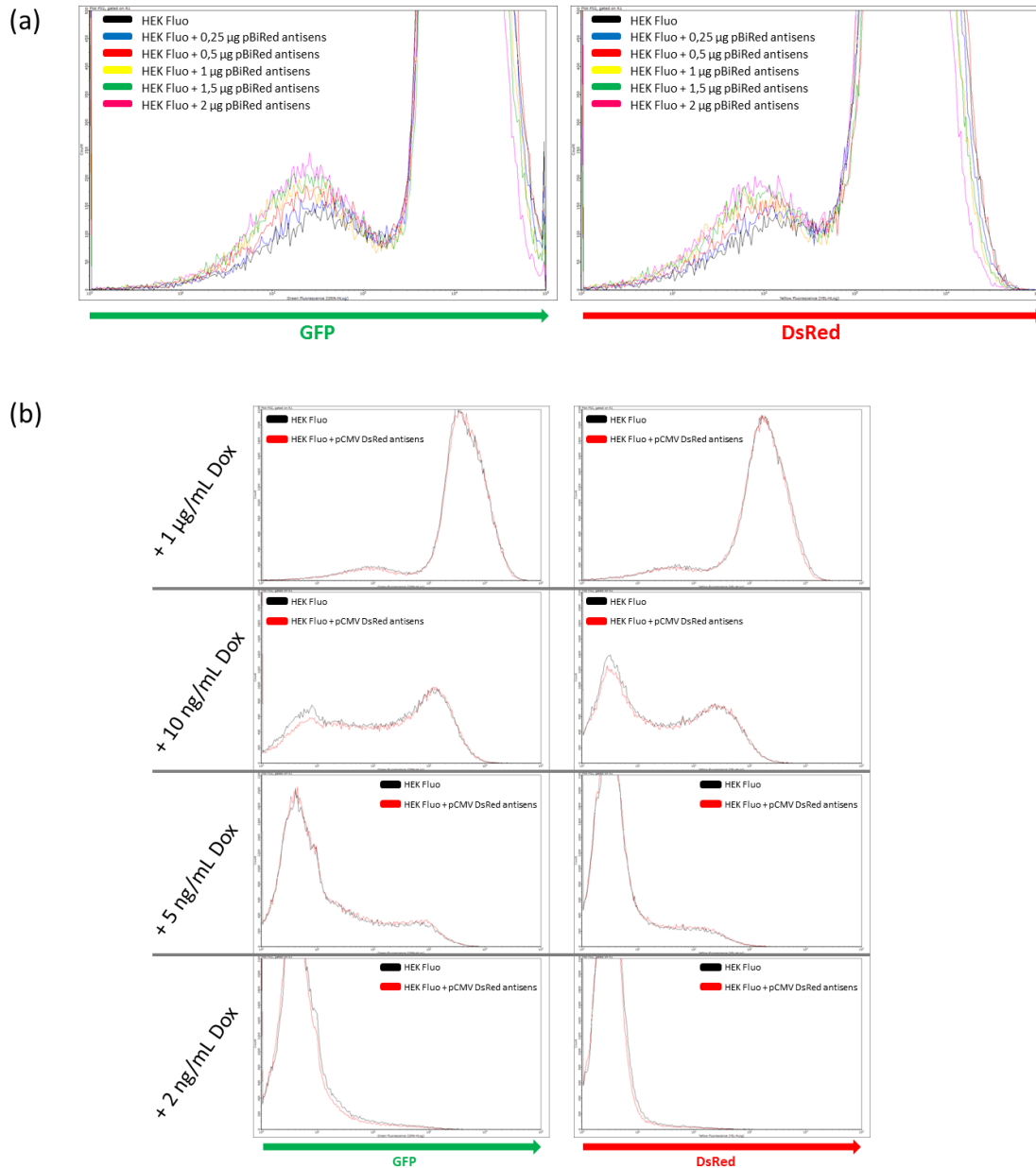
Une autre observation marquante dans cette expérience est que malgré l'absence d'effet significatif sur la cible directe, l'expression de la construction antisens aura un effet aspécifique fort sur la construction pCMV Nir sens utilisée comme contrôle. Ainsi, alors que l'on ne constate aucun silencing sur la construction intégrée par l'expression de pCMV DsRed antisens, on peut remarquer un silencing fort sur l'expression de la construction pCMV Nir sens. Le silencing non spécifique déclenché semble donc similaire à celui des expériences réalisées en transitoire. Ceci est compatible avec la formation du duplex ARN sens issu de l'intégré et antisens issu du transitoire dont on observerait l'activité non spécifique sur le contrôle exprimé en transitoire tandis que l'activité sur la cible directe – qu'elle soit spécifique ou non – reste très limitée.

La différence de réponse entre les constructions pCMV DsRed antisens et pBiRed antisens trouve peut-être son origine dans la force respective des deux promoteurs. En effet, dans notre modèle cellulaire, le promoteur inductible par la tétracycline peut conférer des niveaux d'expression plus forts que le promoteur constitutif CMV. Si le niveau d'expression de l'ARN antisens est un facteur critique de la réponse, l'association d'une construction intégrée (exprimée assez fortement dans toutes les cellules) et d'une construction exprimée en transitoire (dont le niveau n'est comparable à celui du sens que dans une faible proportion des cellules) introduit un différentiel d'expression qui pourrait masquer la réponse recherchée. Pour tester cette hypothèse j'ai réalisé une dose réponse avec une quantité croissante de construction pBiRed antisens, allant de 0, 25 µg de plasmide qui est la dose habituelle dans nos expériences à 8 fois cette dose, soit 2 µg de plasmide (Fig. 28a).

Malgré la transfection d'une dose importante de construction antisens, les résultats ne montrent qu'un effet faible sur l'expression de la construction intégrée au sein de la population principale. De plus cet effet est comparable entre la cible directe (DsRed) et la



GFP suggérant qu'il s'agit d'une activité non spécifique. Ainsi, même en tentant de forcer l'expression il semble difficile d'observer un silencing significatif dû à l'antisens sur l'expression de la construction intégrée.



**Figure 28 : Effet de l'expression relative sens / antisens sur le silencing d'une cible intégrée**

Expériences réalisées dans les cellules HEK Fluo obtenus après intégration de la construction pBiFluo au sein du génome de cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide ou plus pour la dose réponse et analyse 48h après transfection en cytométrie en flux (GUAVA). L'ajout de la doxycycline au milieu de culture permettant l'expression de la construction intégrée est fait le jour de la transfection. (a) Dose réponse avec quantité croissante de plasmide pBiRed antisens transfecté de 0 à 2 µg. Les résultats sont présentés sous forme d'histogrammes de fluorescence GFP et DsRed. (b) Quatre doses de doxycycline sont testées pour modifier l'expression de la construction intégrée (de 2 ng/mL à 1 µg/mL). Ces 4 doses sont testées dans deux conditions soit sans antisens (courbes noires) soit avec la construction pCMV DsRed antisens (courbes rouges). Pour chaque dose de doxycycline, la quantité de construction pCMV DsRed antisens est constante (0.25 µg).

Dans la même logique, la construction intégrée étant inductible à la tétracycline, j'ai exploré la possibilité de jouer sur la concentration de doxycycline pour faire en sorte que l'expression de la construction intégrée soit du même niveau que celle de la construction exprimée en transitoire (Fig. 28b). L'idée étant d'analyser si la différence de niveau d'expression est vraiment la cause de l'absence d'effet. Cette expérience est réalisée à quantité constante de construction pCMV DsRed antisens transfectée. A toutes les doses utilisées l'effet de l'antisens reste très faible, même si l'on s'intéresse aux cellules qui n'expriment la construction cible qu'à un niveau faible.

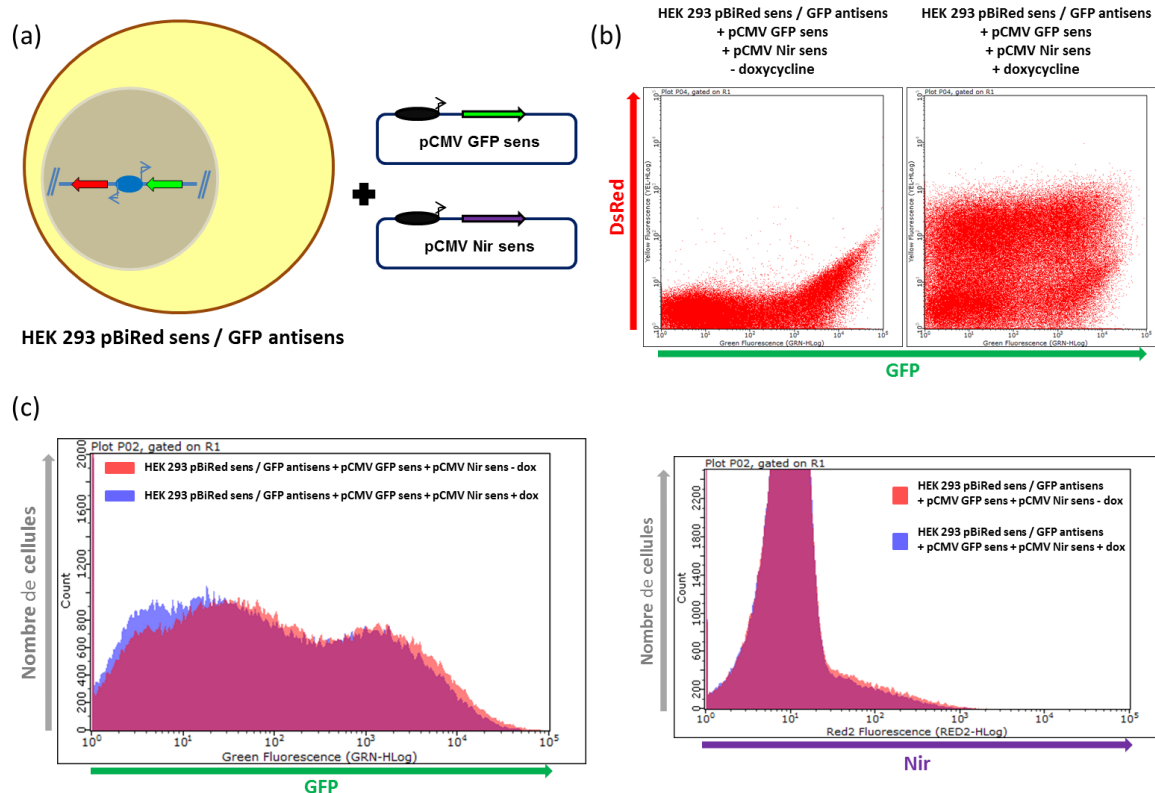
L'expression de l'ARN antisens en transitoire ne semble pas permettre un silencing convaincant, du moins sur une cible intégrée. Cette donnée soulève plusieurs questions car malgré tout, celui-ci permet la mise en place d'un silencing aspécifique, suggérant la formation d'un duplex entre les transcrits sens et antisens. Une possibilité serait que les deux transcrits étant d'origine différente, la transcription ne s'effectue pas dans les mêmes domaines du noyau, ce qui diminuerait la probabilité de la formation des duplex dans le noyau (et donc d'une action spécifique ?) tout en maintenant la possibilité d'une formation dans le cytoplasme au moins dans certaines cellules (et donc d'une réponse non spécifique ?).

Une expérience symétrique a été réalisée où, cette fois, l'antisens est exprimé à partir d'une construction intégrée tandis que sa cible est exprimée en transitoire (Fig. 29). Ici l'antisens est exprimé à partir d'une construction dérivée de pBiFluo où la séquence codante de la GFP est positionnée en orientation antisens. Cette construction permet une co-expression de l'antisens GFP et de la DsRed sous dépendance du même promoteur et donc de suivre l'expression de l'antisens à travers celle de la DsRed. Dans cette expérience, de la même façon que pour l'expression de la construction antisens en transitoire, l'expression de l'ARN antisens n'a qu'un faible effet sur sa cible (silencing < 20%) à la limite de la significativité. Cette donnée confirme bien que la faiblesse de l'effet observé ne vient pas de la différence du niveau d'expression puisqu'ici l'expression de la DsRed est observable dans des cellules qui expriment tous les niveaux de GFP. L'hypothèse d'une absence de silencing liée à une absence d'interaction dans le noyau entre les transcrits issus de loci différents semble donc renforcée. Une deuxième observation importante est que dans cette expérience on n'observe pas la mise en place d'un silencing aspécifique via l'interaction de l'antisens (issu



# Résultats

de la construction intégrée) et du sens (issus de l'expression épisomale). Ainsi, il semblerait que l'origine de l'ARN antisens soit importante pour la mise en place du mécanisme non spécifique. Cela peut être interprété comme un rôle antiviral de ce mécanisme qui va être activé en présence d'un ARN antisens d'origine externe au génome mais qui va tolérer cet ARN antisens s'il est transcrit à partir du génome. Il est donc nécessaire dans ce cas que la cellule soit capable de différencier les deux types de transcrits selon leur origine.



**Figure 29 : Effet d'un antisens intégré sur une construction exprimée en transitoire**

Expériences réalisées dans les cellules HEK pBiRed sens / GFP antisens obtenue après intégration de la construction au sein du génome de cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). L'ajout de la doxycycline pour l'induction de la construction antisens se fait le jour de la transfection. (a) Schéma de l'expérience menée avec la construction pBiRed sens / GFP antisens intégrée et les constructions pCMV GFP sens et pCMV Nir sens exprimées en transitoire. (b) Profils de fluorescence obtenus 48h après transfection. Les deux profils correspondent aux conditions sans et avec doxycycline qui permet l'expression de la construction intégrée. (c) Histogrammes de fluorescence GFP et Nir. On peut y observer pour chacun la comparaison entre la condition sans (Courbe rouge) et avec doxycycline (Courbe bleue).

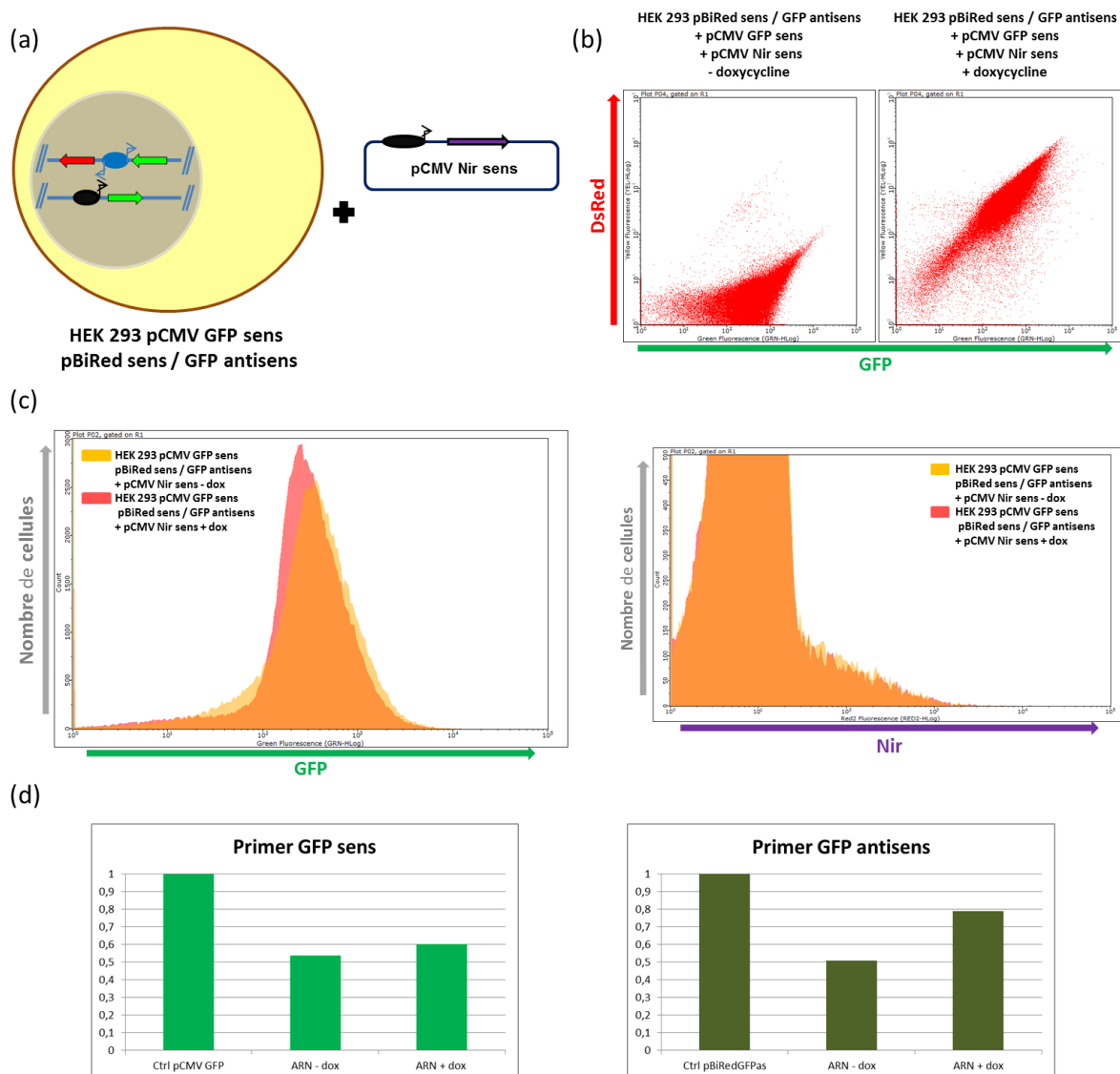
## 2.1.2 – L'effet d'un antisens exprimé en intégré

En l'absence d'un silencing spécifique lorsque l'un des transcrits (sens ou antisens) est exprimé à partir d'un épisome, j'ai exploré la réponse lorsque les deux transcrits sont issus de gènes intégrés. La construction pCMV GFP sens, pour la cible, et la construction pBiRed sens GFP antisens (dérivé de la construction pBiFluo) ont été co-transfectées dans des



# Résultats

cellules HEK 293 Tet-on et des clones stables sélectionnés. Cela donne un modèle de cellules exprimant de manière constitutive la GFP et dans lesquelles l'expression de l'ARN antisens GFP peut-être induite par ajout de la doxycycline au milieu de culture. De plus, l'expression de l'ARN antisens peut être suivie via l'expression de la DsRed placée sous la dépendance du même promoteur.



**Figure 30 : Effet d'un antisens intégré sur une construction cible intégrée**

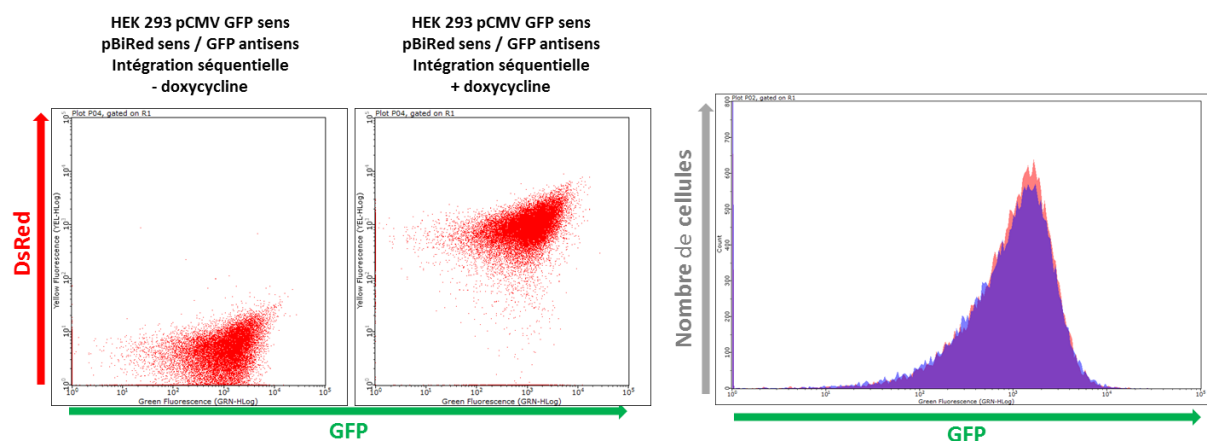
Expériences réalisées dans les cellules HEK pCMV GFP sens pBiRed sens / GFP antisens obtenue après intégration des constructions au sein du génome de cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide pCMV Nir sens comme contrôle de la réponse aspécifique et analyse 48h après transfection en cytométrie en flux (GUAVA). L'ajout de la doxycycline pour l'induction de la construction antisens se fait le jour de la transfection.

(a) Schéma de l'expérience menée avec les constructions pCMV GFP sens et pBiRed sens / GFP antisens intégrées et la construction pCMV Nir sens exprimée en transitoire. (b) Profils de fluorescence obtenus 48h après transfection. Les deux profils correspondent aux conditions sans et avec doxycycline qui permet l'expression de la construction antisens intégrée. (c) Histogrammes de fluorescence de GFP et Nir. On peut y observer pour chacun la comparaison entre la condition sans (courbe jaune) et avec doxycycline (courbe rouge). (d) Histogrammes des niveaux d'expression des transcrits GFP sens et antisens, obtenus par RT-PCR quantitative, en absence ou présence de doxycycline dans le milieu de culture. Les contrôles correspondent aux plasmides en solution utilisés pour l'intégration.



# Résultats

La comparaison des données avec et sans doxycycline est présentée Figure 30. Comme on peut le constater, l'ajout de la doxycycline va permettre l'expression de l'ARN antisens GFP observable à travers l'expression de la DsRed et confirmée par une mesure d'ARN par RTqPCR. Néanmoins, malgré cela l'expression de la cible n'est pas clairement affectée par la présence de l'ARN antisens. Ainsi, l'intégration des deux constructions sens et antisens dans le génome de la cellule ne conduit pas à un silencing spécifique sur la cible directe. Au vu des profils de diffusion obtenus en présence de doxycycline, nous serions plus enclins à envisager un effet inducteur sur la cible par l'expression de l'ARN antisens. En effet, il est étonnant de voir à quel point les fluorescences GFP et DsRed sont corrélées. On retrouve ainsi que les forts expresseurs de GFP sont également les forts expresseurs de DsRed alors que les deux protéines s'expriment à partir des deux constructions différentes avec des promoteurs différents. Ces données montrent une complexité supplémentaire du modèle où vraisemblablement l'induction de l'activité du promoteur pBi aura un effet enhancer sur le promoteur CMV. Il est plausible que cet effet provient de la stratégie de co-transfection qui a été utilisée pour obtenir ces clones stables, ce qui crée la possibilité d'intégrations en tandem.



**Figure 31 : Effet d'un antisens intégré sur une construction cible intégrée. Intégration séquentielle**

Expériences réalisées dans les cellules HEK pCMV GFP sens pBiRed sens / GFP antisens obtenues après intégration des constructions au sein du génome de cellules HEK 293 3G Tet-on de façon séquentielle. Analyse à 48h en cytométrie en flux (GUAVA) après induction de l'expression de l'antisens ou non par ajout de doxycycline au milieu de culture. Les deux profils de fluorescence et l'histogramme de fluorescence GFP montrent la comparaison entre les deux conditions.

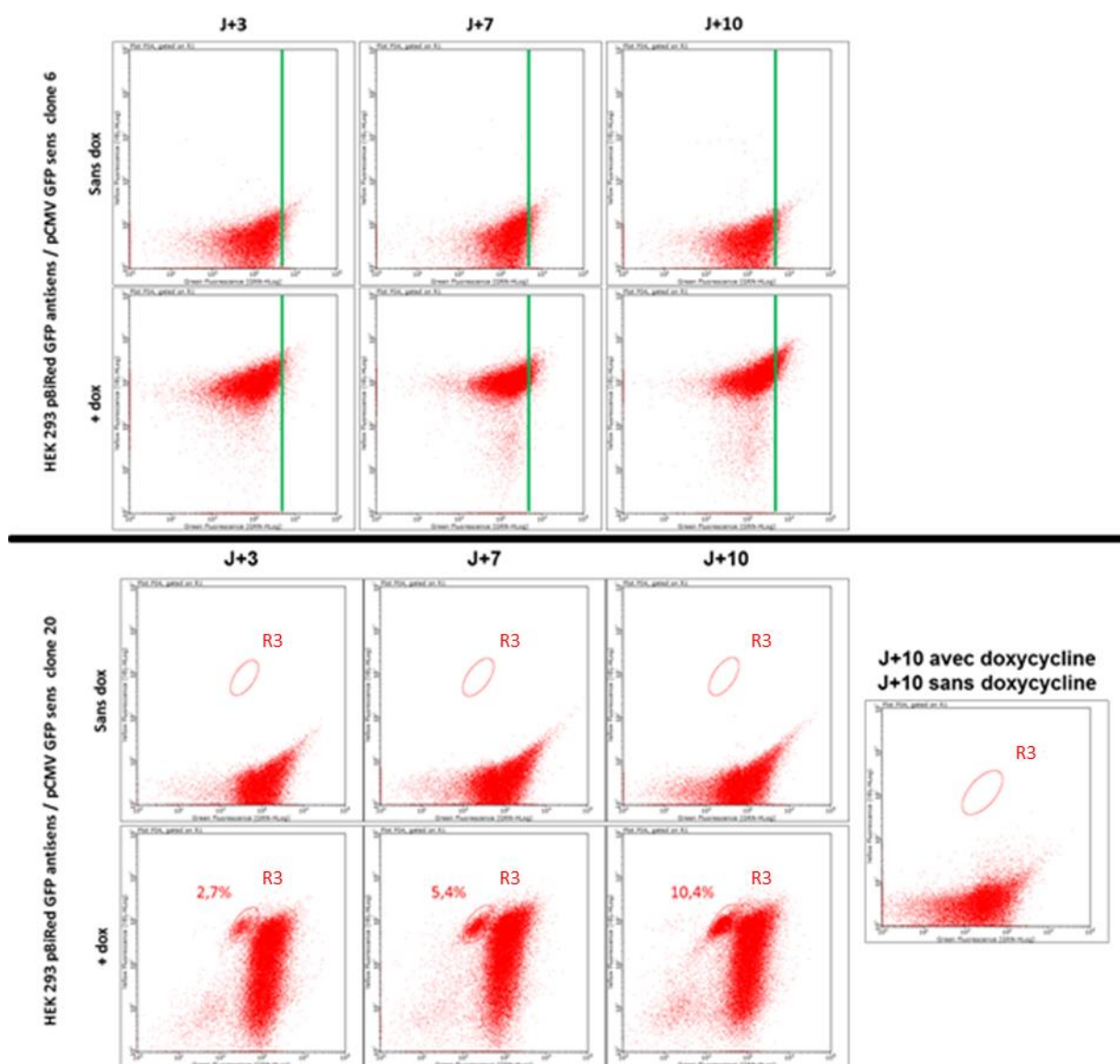
Pour essayer de diminuer les effets enhancer liés à l'induction de l'expression de l'antisens, des clones stables ont été réalisées avec intégration des constructions de façon séquentielles. Les données présentées figure 31 montrent, qu'après induction, l'expression





# Résultats

de la GFP et celle de la DsRed sont nettement moins corrélées que dans la Figure 30. Ainsi une transfection séquentielle permet de réduire l'effet enhanceur lié au promoteur pBi. Cependant, malgré cela, l'expression de l'ARN antisens ne permet pas la mise en place d'un silencing. On ne distingue aucune différence entre le clone exprimant l'antisens et le clone ne l'exprimant pas au niveau de l'intensité de fluorescence GFP.



**Figure 32 : Cinétique longue pour les clones stables intégrés**

Expériences réalisées sur deux clones stables HEK 293 pCMV GFP sens pBiRed sens / GFP antisens. Pour l'un obtenu par intégration séquentielle (clone 6) et l'autre intégration simultanée (clone 20). Analyse en cytométrie (GUAVA) après 3 jours, 7 jours et 10 jours avec ou sans doxycycline dans le milieu de culture. Dans le cas du clone 20, une analyse supplémentaire sur 10 jours sans doxycycline après 10 jours avec doxycycline a été faite pour rechercher un maintien du silencing observé.



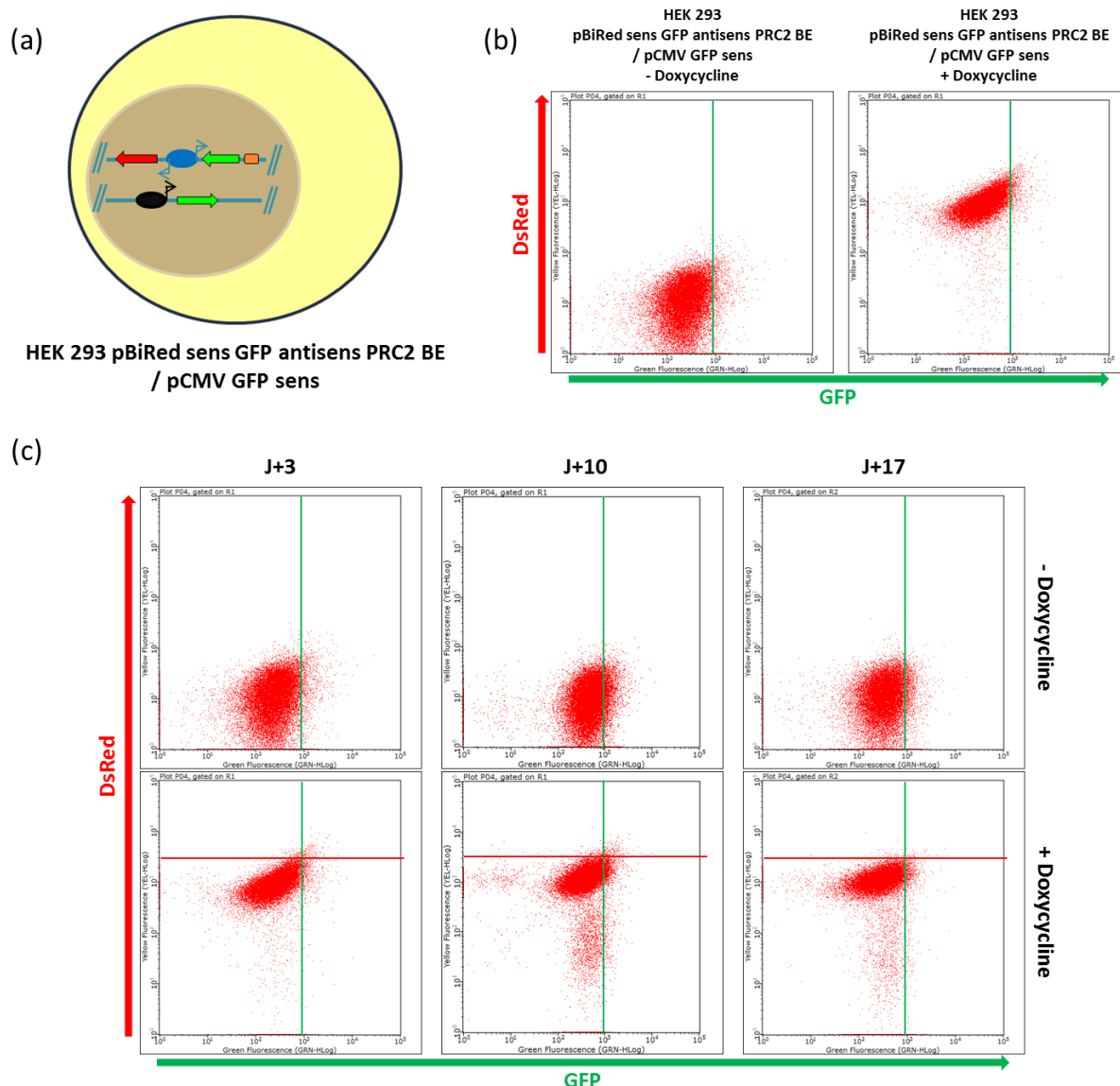
Nos expériences en clones stables sont jusqu'ici uniquement faits sur des cinétiques courtes, la lecture des cellules au cytomètre en flux ayant lieu 48 heures après la transfection et/ou l'induction de l'expression de l'antisens. Cette cinétique est suffisante si l'on cherche à observer des mécanismes de régulation post-transcriptionnelle mais pourrait être trop courte pour un silencing impliquant un changement d'organisation de la chromatine. A cette fin, j'ai suivi l'évolution de l'expression des gènes rapporteurs en présence de l'ARN antisens sur une période plus longue (10 jours dans le cas présenté Figure 32). Le maintien sur plusieurs jours de ces cellules en milieu de culture avec doxycycline n'a pas d'impact marqué sur le profil d'expression. On peut cependant dans certains cas comme le clone 20 de la figure 32, observer l'apparition d'une sous-population semblant être mise sous silence de façon spécifique (diminution de la GFP en présence d'une forte expression de la DsRed). Cette population est identifiée par la région R3 sur la Figure 32 et la proportion de cellules soumises à ce silencing augmente progressivement au cours du temps passant de 2,7 à 10,4 % de J + 3 à J + 10. L'arrêt de l'induction par la doxycycline ne permet pas de suivre une population dans laquelle ce silencing perdurerait au-delà de l'expression de la DsRed. Le clone 20 étant un clone réalisé par intégration simultanée des deux constructions, on peut se demander si ceci ne joue pas un rôle dans le silencing observé, puisqu'il n'a pas été retrouvé dans les clones réalisés en intégration séquentielle. La possibilité d'un effet en *cis* de l'antisens sera étudiée plus loin.

Il se dégage de ces résultats que l'expression de l'antisens n'a pas d'effet significatif sur sa cible directe et cela même si les deux constructions sont intégrées. Il semblerait que dans les conditions d'une expression en *trans* d'une construction antisens, la simple complémentarité de séquence ne soit pas suffisante pour la mise en place d'un silencing. Il est possible que l'ajout de séquence spécifique permettant le recrutement de partenaires protéiques soit nécessaire comme c'est le cas dans pour les modèles *in vivo* tels que HOTAIR, ANRIL ou l'ARN Xist.



## 2.1.3 – Effet de l'ajout d'une séquence de recrutement

Les expériences précédentes montrent qu'à elle seule la transcription antisens n'a qu'un effet limité sur l'expression du gène « sens ». Comme il a été proposé, et validé expérimentalement au moins dans certains cas, que l'action des antisens pouvait passer par le recrutement de complexes de modification de la chromatine, j'ai donc modifié la construction pStan pour tester cette hypothèse dans le cadre de ce modèle.



**Figure 33 : Test du clone stable pBiRed sens GFP antisens PRC2 BE / pCMV GFP**

Le clone stable est réalisé après intégration séquentielle des constructions pCMV GFP sens et pBiRed sens GFP antisens PRC2 BE en cellule HEK 293 Tet-on. (a) Représentation schématique du clone utilisé. Les gènes sont présentés sous forme de flèches (Rouge = DsRed ; Vert = GFP). Le « PRC2 Binding Element » est présenté sous la forme d'une cassette orange. (b) Profils de fluorescence obtenus 48h après induction ou non par la doxycycline de la construction antisens en cytométrie en flux (GUAVA). (c) Profils de diffusion des cinétiques longues (J+3 à J+17) en présence ou absence de doxycycline obtenus par analyse en cytométrie de flux (GUAVA). Les barres vertes et rouges sont positionnées au même niveau sur l'ensemble des profils.

Notre choix s'est porté sur la séquence minimale du « PRC2 Binding Element » présente dans le long ARN non codant HOTAIR (Wu *et al*, 2013). Il a été proposé que cette séquence permettait le recrutement du complexe Polycomb via une structure secondaire en tige boucle. Notre hypothèse est qu'en insérant cet élément dans notre séquence antisens, le recrutement du complexe régulateur Polycomb PRC2 sur le transcrit naissant pourra induire un silencing transcriptionnel du gène sens.

Pour tester cette hypothèse, une construction « pBiRed sens GFP antisens PRC2 BE » a été réalisée et des clones stables isolés dans les cellules HEK 293 Tet-on (Fig. 33a). Pour éviter les effets de transactivation, l'intégration avec la cible pCMV GFP sens a été faite de manière séquentielle. Malgré la présence de cette séquence de fixation du complexe Polycomb, l'expression de l'antisens ne permet toujours pas un silencing significatif sur la cible GFP (Fig. 33b). L'ajout au milieu de culture de la doxycycline qui induit l'expression de l'ARN antisens ne modifie en rien l'expression de la GFP. Comme pour la majorité des expériences, les analyses en cytométrie en flux sont faites à 48 heures après l'induction par la doxycycline. Une expérience avec une cinétique plus longue a été réalisée (Fig. 33c). Cependant, même après 17 jours de culture en présence de doxycycline, l'effet de l'antisens n'est pas observable. Ces données laissent à penser que l'action en *trans* d'un ARN antisens est difficile à mettre en place et nécessite d'autres paramètres.

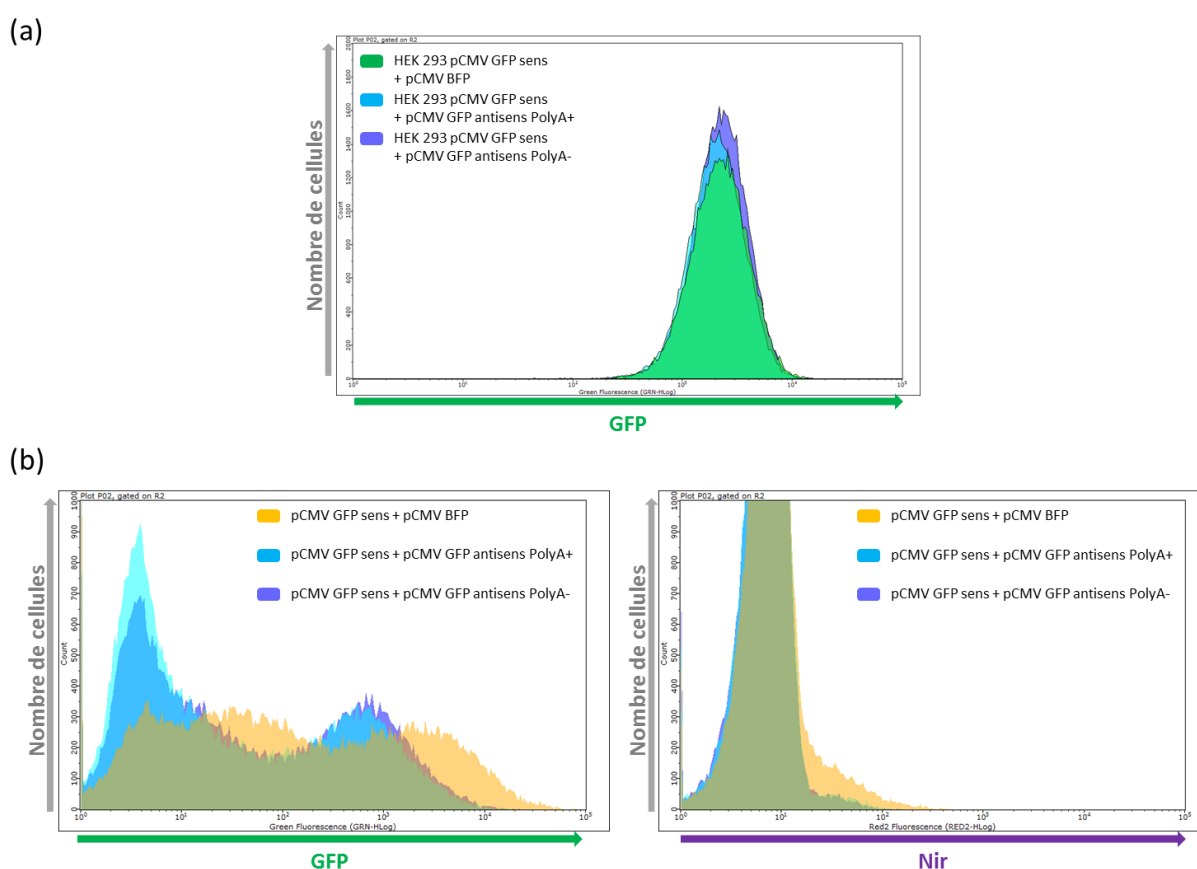
## 2.1.4 – Effet de l'absence de séquence de polyadénylation

Une des différences entre notre ARN antisens et la plupart des situations *in vivo* est que celui-ci est d'une taille modeste et d'une faible complexité. En conséquence, il se pourrait qu'il soit transcrit et exporté dans le cytoplasme rapidement et ne reste pas suffisamment longtemps à proximité de son site de transcription pour permettre une action sur la chromatine. De plus, une telle cinétique pourrait biaiser la réponse cellulaire en favorisant la formation d'ARN double brin dans le cytoplasme. Des travaux menés par l'équipe de Nick Proudfoot ont montré la possibilité d'induire un silencing sur une cible intégrée par la co-transfection d'un plasmide sens et antisens, donc via une action en *trans* (Gullerova and Proudfoot, 2012). Leurs données montrent, entre autres, que si la formation du duplex d'ARN s'effectue dans le noyau il n'y a pas le déclenchement de réponse aspécifique de type interféron. De plus, dans cette étude, ils utilisent des vecteurs d'expression dépourvus de



séquence de polyadénylation ce qui pourrait favoriser la persistance des transcrits dans le noyau.

J'ai tout d'abord testé l'activité d'une construction pCMV GFP antisens dans laquelle le site de polyadénylation (construction « PolyA - ») sur le clone stable HEK 293 pCMV GFP sens a été délété. En parallèle, j'ai transfecté la construction initiale (pCMV GFP antisens « polyA+ ») (Fig. 34a). Cependant, l'antisens PolyA- n'a pas d'effet plus important sur la cible intégrée que la construction antisens initiale.



**Figure 34 : Test de l'effet de l'absence de site de polyadénylation sur le silencing antisens**

(a) Expérience réalisée dans le clone stable HEK pCMV GFP sens. Transfection de 0,25 µg de chaque plasmide. Analyse à 48h au cytomètre de flux après transfection. Résultats présentés sous forme d'histogrammes de fluorescence GFP. Expérience faite en trois conditions, en absence d'antisens avec le plasmide contrôle pCMV BFP sens (courbe verte), avec la construction antisens contenant la séquence PolyA (courbe bleue) ou bien avec la construction antisens sans la séquence PolyA (courbe violette). (b) Expérience de transfection transitoire. Analyse à 48h au cytomètre de flux après transfection au phosphate de calcium. Transfection de 0,25 µg de chaque plasmide. Résultats représentés sous forme d'histogrammes de fluorescence GFP et Nir. Expérience faite en trois conditions, en absence d'antisens avec le plasmide contrôle pCMV BFP sens (courbe jaune), avec la construction antisens contenant la séquence PolyA (courbe bleue) ou bien avec la construction antisens sans la séquence PolyA (courbe violette).

Puisque la logique de cette approche est de biaiser la rétention nucléaire du transcrit antisens j'ai aussi recherché si, en transitoire, l'ARN antisens PolyA<sup>-</sup> n'induisait pas de réponse aspécifique (Fig. 34b). Mais là encore, l'ARN antisens PolyA<sup>-</sup> montre un profil similaire à l'ARN antisens PolyA<sup>+</sup> avec un silencing aspécifique observable sur l'expression de pCMV Nir sens.

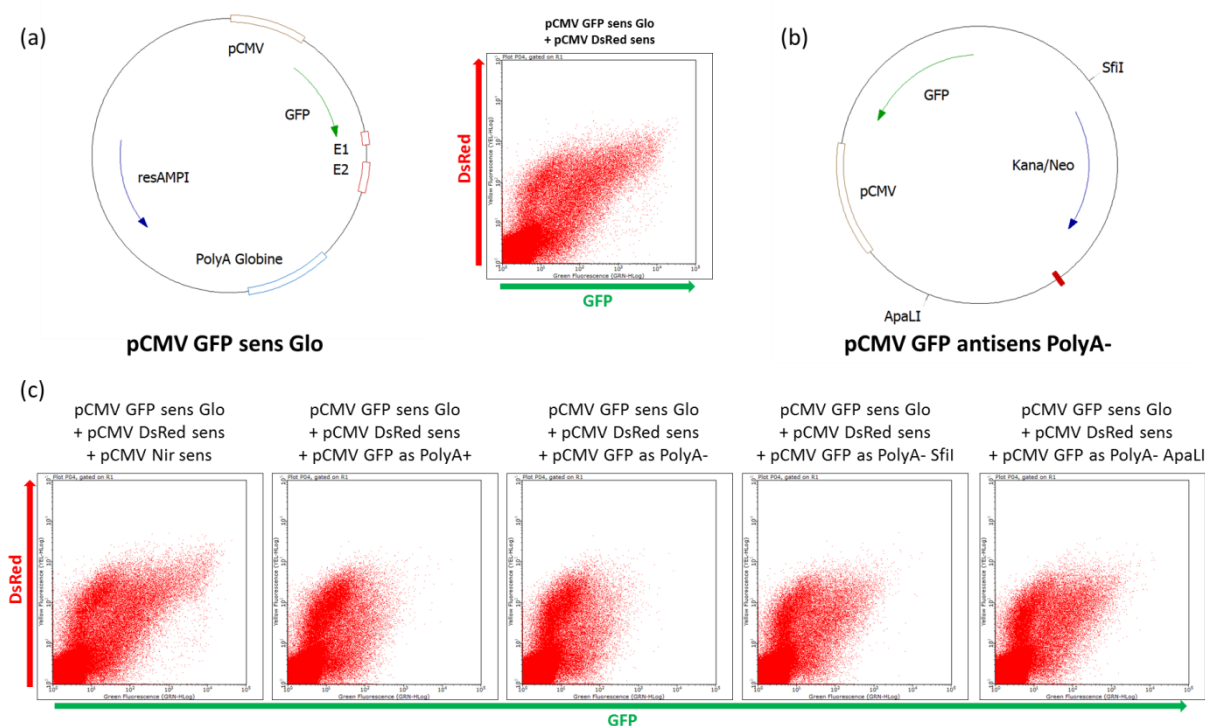
Puisque l'absence de site de polyadénylation ne suffit pas à induire un silencing spécifique, il est possible que la structure du transcrit cible soit aussi en cause. De fait, le même commentaire, petite taille et faible complexité, s'applique aussi bien aux aspects sens qu'antisens de nos constructions. C'est pourquoi j'ai aussi utilisé comme cible la construction pCMV GFP sens Glo. Cette construction a la particularité d'ajouter à la séquence GFP les deux premiers introns et exons de la  $\beta$ -globine (Fig. 35a). Cela augmente la taille du transcrit GFP et introduit une étape de maturation qui devrait conduire à une présence plus longue du transcrit GFP dans le noyau. Les propriétés du transcrit cible sont un paramètre peu discuté jusqu'à présent mais potentiellement aussi important que celles du transcrit antisens. Dans notre cas, on peut constater que la présence des exons et introns a un effet sur l'expression de la construction dont l'expression est plus faible qu'une construction pCMV GFP sens standard (Fig. 35a). Ce résultat est inattendu compte tenu du lien général entre épissage, transport et traduction des ARN messagers, mais reflète probablement la synthèse et l'activité d'une protéine de fusion GFP- $\beta$ -globine ; une telle organisation de la construction étant indispensable si l'on veut éviter une régulation par la voie NMD (Nonsense Mediated mRNA Decay). L'ajout des constructions antisens a donné le même type de réponse entre les ARN antisens PolyA<sup>+</sup> et PolyA<sup>-</sup> (Fig. 35a). A la vue de ces résultats, on distingue plusieurs types de réponse en fonction des sous-populations concernées. On observe un fort silencing de la GFP qui affecte peu une population pure DsRed déjà présente dans le profil sans antisens. La présence de cette sous-population gêne quelque peu l'analyse car donne l'impression d'un silencing spécifique GFP plus que d'un silencing non spécifique attendu dans ce cas de figure.

Les données montrent que, globalement, l'absence de site de polyadénylation ne permet pas d'induire un silencing spécifique. Cependant, il semble que malgré l'absence de PolyA, l'ARN antisens soit capable d'être exporté dans le cytoplasme. Cette observation est surprenante, car en l'absence d'une voie de transport alternative (comme par exemple dans



# Résultats

le cas des ARNm des histones) on s'attendrait à ce que les transcrits polyA- ne soient pas exportés. J'ai donc recherché la présence de séquences potentielles de polyadénylation dans les constructions antisens PolyA- (Fig. 35b). Comme on peut le remarquer, deux séquences potentielles de polyadénylation sont présentes dans le corps du plasmide. Pour éliminer l'éventualité que ces séquences puissent jouer un rôle, une linéarisation du plasmide a été faite en amont de ces séquences, avant la transfection (Fig. 35c). Mais, même si on peut constater une diminution du silencing non spécifique (visible sur la DsRed), ce dernier est toujours présent. De plus, un changement similaire est observable sur la cible directe (GFP), ce qui indique que l'on n'est toujours pas parvenu à renforcer un silencing spécifique. Aussi, une interprétation plausible de ces résultats est que l'effet majeur de la linéarisation est de diminuer l'expression de la construction antisens mais sans changer le mécanisme sous-jacent.



**Figure 35 : Test de l'effet d'un transcrit antisens PolyA- sur un transcrit épissé**

(a) Schéma de la construction pCMV GFP sens Glo et profil de fluorescence en présence de pCMV DsRed sens. Profil obtenu par cytométrie en flux 48h après transfection, transfection de 0,25 µg de chaque plasmide. (b) Schéma de la construction pCMV GFP antisens PolyA-. Les deux sites SfiI et ApaLI utilisés pour sa linéarisation et également la position des deux sites de polyadénylation sont indiqués (marqueurs rouges). (c) Profils de fluorescence obtenus par analyse en cytométrie en flux 48h après transfection, transfection de 0,25 µg de chaque plasmide. Différentes conditions sont testées en absence ou en présence de la construction pCMV GFP antisens linéarisée ou non.

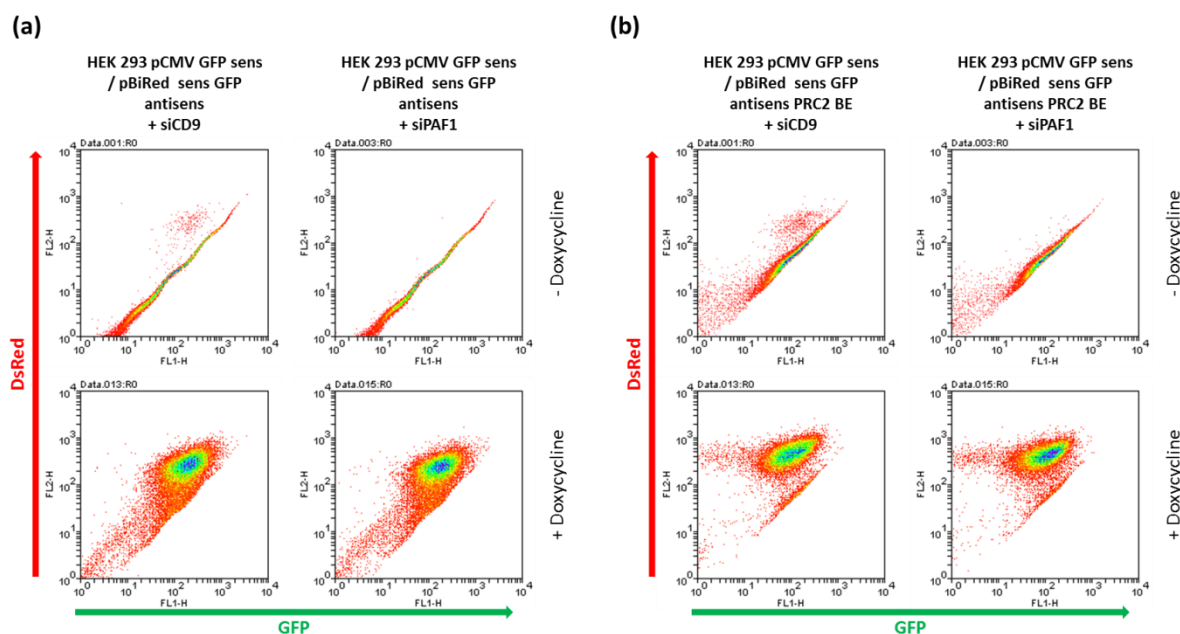
## 2.1.5 – Effet de l'inhibition de PAF1

Toujours dans l'optique d'augmenter le temps de résidence nucléaire des transcrits, l'inhibition de la protéine PAF1 a été testée. PAF1 est une protéine associée à la polymérase ARN comme son nom l'indique, « RNA-Polymérase associated Factor 1 ». Elle participe, en autres, à l'élongation de la transcription. Chez *S. pombe*, des travaux sur le silencing transcriptionnel médié par les petits ARN ont montré que l'inhibition de PAF1 avait pour effet de rendre accessible le transcrit naissant en cours d'élongation. L'absence de PAF1 chez *S.pombe* permet ainsi de mettre en place un silencing transcriptionnel en réponse à des petits ARN double brin (Kowalik *et al*, 2015). J'ai donc testé si l'inhibition de PAF1 dans les cellules HEK293 était capable de rendre le transcrit sens naissant accessible à un long ARN antisens. Ce test a été effectué en clone stable HEK 293 pCMV GFP sens / pBiRed sens GFP antisens avec ou sans la présence de l'élément de recrutement du complexe Polycomb PRC2. Comme on le montre la figure 36, aucun effet de l'inhibition de PAF1 par un siARN n'est observable. La même réponse est obtenue avec la construction contenant le site de recrutement du complexe Polycomb PRC2.

Ainsi l'inhibition de la protéine PAF1 ne permet pas la mise en place d'un silencing dans notre modèle d'ARN antisens exprimé en *trans*. L'hypothèse la plus vraisemblable est que les deux transcrits sens et antisens ne se rencontrent pas dans le noyau mais dans le cytoplasme, ne permettant pas un mécanisme transcriptionnel. C'est pourquoi l'étude du modèle en *cis* a été poursuivie.

En conclusion, malgré plusieurs tentatives nous n'avons pas réussi à observer la mise en place d'un silencing par un ARN antisens exprimé en *trans*. Une des possibilités évoquées dans les travaux sur les ARN antisens est que la probabilité de rencontre entre les ARNs sens et antisens soit en général trop faible lorsqu'ils sont transcrits à partir de loci distincts. Leur rencontre restant cependant possible dans le cytoplasme du fait de leur temps de résidence mais ne conduisant pas à un silencing spécifique.





**Figure 36 : Test de l'effet du siPAF1 sur le silencing en *trans* par un ARN antisens intégré**

(a) Profils de fluorescence GFP/DsRed obtenus par cytométrie en flux (FACS Calibur) 48h après transfection des cellules HEK 293 pCMV GFP sens / pBiRed sens GFP antisens par le siARN contrôle (siCD9) ou le siPAF1 à une concentration de 5 nM. L'ajout ou non de la doxycycline se fait 24h après la transfection. (b) Profils de fluorescence GFP/DsRed obtenus par cytométrie en flux (FACS Calibur) à 48h après transfection des cellules HEK 293 pCMV GFP sens / pBiRed sens GFP antisens PRC2 BE par le siARN contrôle (siCD9) où le siPAF1 a une concentration de 5 nM. L'ajout ou non de la doxycycline se fait 24h après la transfection.

## 2.2 –Silencing en *cis* d'une construction intégrée

L'absence d'un silencing spécifique induit par un ARN antisens exprimé en *trans* et l'observation que, chez les mammifères, les effets sur les gènes cellulaires sont en général attribués à l'expression d'un antisens en *cis* nous ont donc conduit à nous focaliser sur la construction pStan après intégration dans le génome cellulaire.

### 2.2.1 – Analyse du silencing dans la construction pStan intégrée

Comme pour l'ensemble des clones stables précédemment réalisés, la construction pStan est intégrée dans les cellules HEK 293 Tet-on. Cette construction, déjà décrite dans la partie transfection transitoire, permet une expression constitutive de la protéine GFP. C'est l'ajout de la doxycycline au milieu de culture qui permet l'induction du promoteur pBi qui va alors transcrire un ARN « DsRed sens / GFP antisens ».

Les données sur l'expression de pStan dans les clones stables sont similaires à celles observées en transitoire (Fig. 37a). L'ajout de la doxycycline donne naissance à un silencing

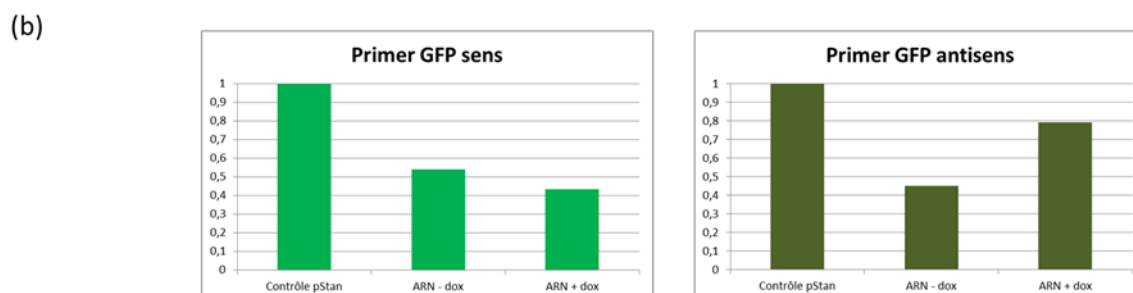
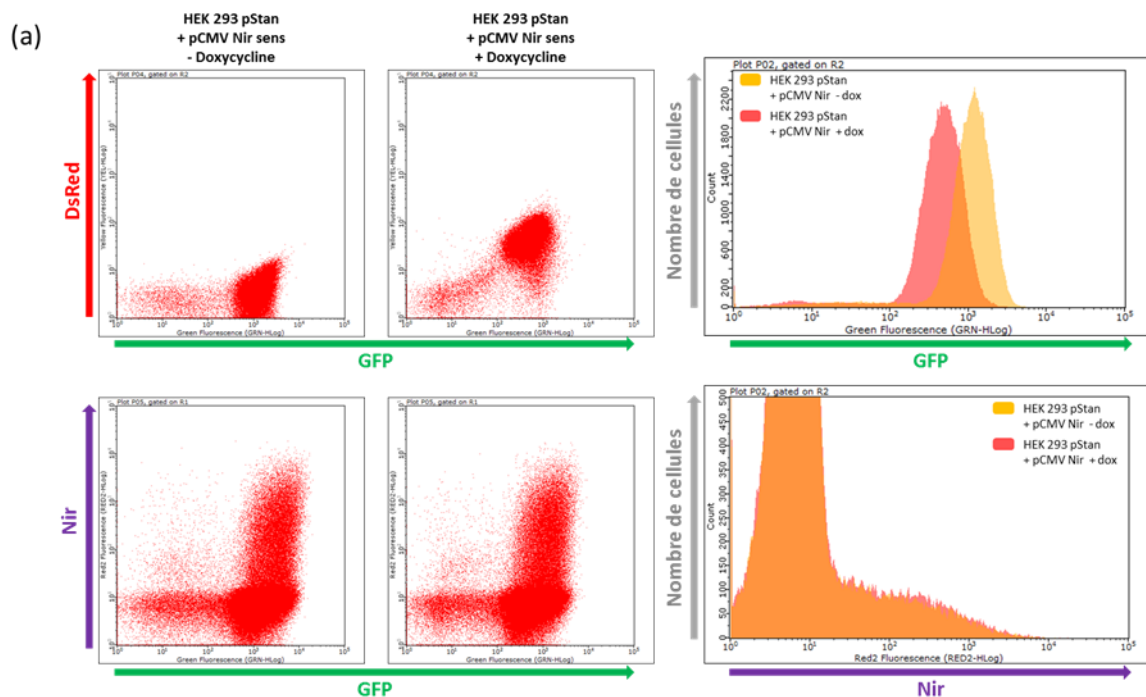


significatif sur la fluorescence GFP (64%). Ici l'expression de la GFP étant plus homogène que lors des expressions en transitoire, il est plus facile de suivre le silencing à travers le déplacement de l'ensemble de la population. La mise en place du silencing coïncide bien avec l'expression de l'ARN antisens comme l'atteste les données de RT-qPCR (Fig. 37b).

Contrairement aux effets de la construction en transitoire et en accord avec les résultats obtenues lors des expériences des constructions antisens intégrées dans un contexte *trans*, l'induction de l'expression de l'antisens par la doxycycline n'affecte pas l'expression du contrôle pCMV Nir sens s'exprimant en transitoire. Pour la première fois, dans ce contexte de l'expression d'un ARN antisens en *cis*, il semble possible de mettre en place un silencing spécifique sur la cible intégrée.

Les données présentées ici sont celles du clone 21 HEK 293 pStan. Cependant, d'autres clones ont été analysés. Le clone 9 montre un silencing plus faible que le clone 21 mais aussi une induction plus limitée de l'expression de l'antisens (mesurée à travers l'expression de la DsRed) (Fig. 37c). Par ailleurs, le clone 23 montre une légère induction de l'expression de la GFP en réponse à la doxycycline mais par d'expression détectable de la DsRed. Ainsi, l'étude de ces quelques clones est en faveur d'un silencing lié à l'expression en *trans* de l'ARN antisens.





(c)

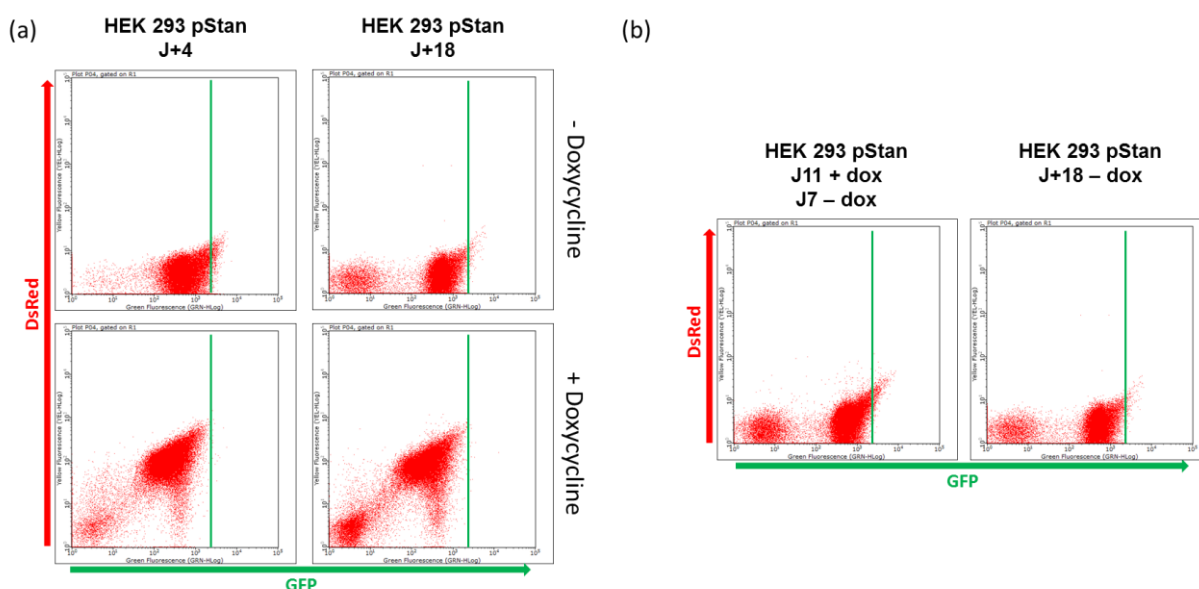
Clone	Effet observé sur la GFP de l'induction de l'antisens	Expression de la DsRed après induction
HEK 293 pStan clone 9	-39%	Faible
HEK 293 pStan clone 21	-50%	Forte
HEK 293 pStan clone 23	+25%	Aucune

Figure 37 : Etude du silencing non spécifique par la construction pStan intégrée.

(a) Profils de fluorescence et histogrammes de fluorescence obtenus par analyse au cytomètre de flux (GUAVA) à 48h des cellules HEK 293 Tet-on pStan clone 21 transfectées par 0,25  $\mu$ g de pCMV Nir sens en présence ou absence de doxycycline. Ajout de la doxycycline à 1  $\mu$ g/mL le jour de la transfection. (b) Histogrammes représentant l'expression relative de l'ARN GFP sens et antisens en présence ou absence de doxycycline. Données obtenues par RT-qPCR avec deux couples de primers et une reverse transcription spécifique pour les transcrits issus du promoteur CMV et du promoteur pBi. Le contrôle utilisé est le plasmide pStan. (c) Tableau récapitulatif des différents clones HEK 293 pStan réalisés et leurs effets.

# Résultats

Pour préciser la nature du silencing, une expérience à plus long terme a été réalisée. Cette cinétique a été réalisée en culture de cellules en présence de doxycycline pendant plus de deux semaines. Le silencing de la GFP n'augmente pas au-delà de celui observé à 48 heures qui semble avoir déjà atteint son maximum (Fig. 38a). Pour tester si ce silencing implique des phénomènes épigénétiques, les cellules sont cultivées pendant 7 jours en l'absence de doxycycline après y avoir été exposées pendant 11 jours (Fig. 38b). On n'observe aucune mémoire du silencing, et dans des observations intermédiaires, il apparaît que le silencing est directement lié à l'expression de l'ARN antisens et diminue en relation avec l'activité du promoteur pBi. Le silencing induit dans ce clone n'est pas transmissible indépendamment de l'expression de l'ARN antisens.



**Figure 38 : Cinétique longue en condition avec doxycycline du clone stable pStan**

(a) Profils de fluorescence obtenus en cytométrie après mise en culture de la lignée HEK 293 Tet-on pStan clone 21 avec ou sans doxycycline durant 4 ou 18 jours. (b) Profils de fluorescence obtenus en cytométrie en flux après mise en culture de la lignée HEK 293 Tet-on pStan en milieu avec doxycycline pendant 11 jours puis passage dans un milieu sans doxycycline durant 7 jours. La comparaison est faite avec la même lignée cellulaire mise en culture dans un milieu sans doxycycline durant 18 jours. Les barres verticales vertes représentent la position initiale de la population principale en terme d'intensité maximale de fluorescence GFP.

En plus de l'évolution de l'expression de la population principale dans les nuages de points de la Figure 38, on peut aussi constater l'apparition d'une petite population qui n'exprime ni la GFP ni la DsRed. Mais comme ce phénomène ne dépend pas de l'exposition à la doxycycline il est plus probable qu'il corresponde à la mise en place au cours du maintien en culture d'un silencing au niveau du site d'intégration dans ce clone stable. La « dérive » de

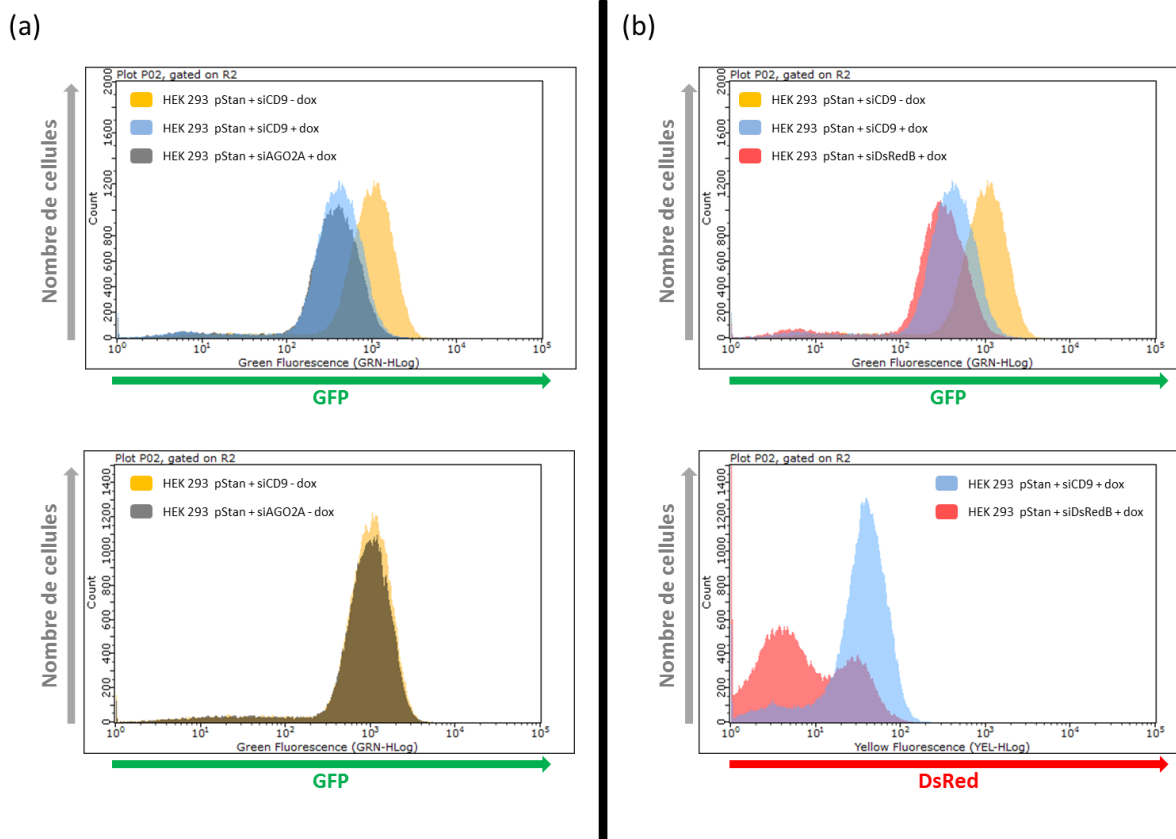
l'expression des clones « stables » est un phénomène fréquemment observé et qui est généralement attribué à la stabilisation de l'environnement chromatinien au niveau des sites d'intégration après la phase initiale au cours de laquelle le gène nouvellement intégré est dans une configuration active.

En ce qui concerne le silencing de la population principale, une première classe de mécanisme est celle de l'interférence par l'ARN, qu'elle agisse aux niveaux post-transcriptionnel ou transcriptionnel. Pour tester cette hypothèse j'ai recherché si ce phénomène était sensible à un silencing d'Ago2 par la transfection de siARN Ago2 (Fig. 39a). Même si le silencing ainsi obtenu n'est pas complet, il conduit à une très forte diminution de l'interférence par l'ARN du moins dans sa forme classique d'hydrolyse post-transcriptionnelle. Typiquement, la réduction du silencing en mode siARN sur une transfection contrôle menée en parallèle est de l'ordre de 75%. Ici nous n'avons observé aucun effet du silencing d'Ago2 sur la diminution d'expression de la GFP. Le siARN ciblant le gène CD9 murin est un contrôle qui a fréquemment été utilisé dans le laboratoire du fait de son absence d'effet sur la croissance cellulaire et de sa capacité à entrer en compétition avec d'autres siARN suggérant qu'il est efficacement incorporé dans les complexes RISC. Les données obtenues avec le siAGO2 semblent donc exclure une implication de l'interférence par l'ARN interférence dans le silencing induit par l'ARN antisens.

Dans la continuité de cette première analyse, j'ai exploré le rôle de l'ARN antisens dans le silencing observé avec cette construction intégrée. Ceci a été entrepris en ciblant le transcrit antisens par un ARN interférant. Le laboratoire dispose d'un siARN très efficace dirigé contre la région codante de la DsRed (diminution de l'expression protéique > 90 %) et qui est donc capable d'empêcher l'accumulation du transcrit antisens (Figure 39b). Bien que dans ces expériences on n'observe plus l'expression de la DsRed, le silencing de la GFP est toujours comparable à celui observé dans le contrôle siCD9. De fait, il est même légèrement supérieur en présence de siDsRedB. On peut donc conclure que l'accumulation de l'ARN antisens dans le cytoplasme, telle qu'elle est mesurée à travers le niveau d'expression de la DsRed, n'est pas un déterminant fort du silencing. *A contrario*, cette expérience ne teste qu'indirectement les activités nucléaires de l'ARN antisens car l'effet du siDsRedB sur les ARN nucléaires est plus modeste que dans le cytoplasme ainsi qu'il a été observé dans le laboratoire par FeiFei Liang. Néanmoins, l'association des données siAgo2 et siDsRedB n'est



pas en faveur de l'implication d'un mécanisme de type interférence par l'ARN, même transcriptionnel, dans le silencing induit par l'antisens.



**Figure 39 : Effet d'un siAGO2 et d'un siDsRed sur le clone stable pStan**

Histogrammes de fluorescence GFP obtenus en cytométrie en flux (GUAVA) 48h après transfection des siARN (5 nM) dans les cellules HEK 293 Tet-on pStan. L'ajout de la doxycycline est faite ou non le jour de la transfection (1 µg/mL). (a) Test de l'effet du siAGO2 sur le silencing. Trois conditions sont présentées. Les cellules HEK 293 Tet-on pStan en absence de doxycycline mais transfectées avec le siARN contrôle siCD9 (courbe jaune), les mêmes cellules transfectées avec le même siCD9 mais, cette fois-ci, en présence de doxycycline (courbe bleue) et toujours les mêmes cellules en présence de doxycycline mais transfectées par le siAGO2 (courbe noire). Le contrôle de l'effet du siAGO2 en absence de l'antisens est également présenté (courbe noire) en comparaison avec les cellules transfectées par le siCD9 contrôle (courbe jaune). (b) Test de l'effet du siDsRedB. Trois conditions sont présentées. Les cellules HEK 293 Tet-on pStan en absence de doxycycline mais transfectées avec le siARN contrôle siCD9 (courbe jaune), les mêmes cellules transfectées avec le même siCD9 mais cette fois-ci en présence de doxycycline (courbe bleue) et toujours les mêmes cellules en présence de doxycycline mais transfectées par le siDsRedB (courbe rouge). Le test de l'efficacité du siDsRedB est également présenté dans l'histogramme du bas, le nombre de cellules en fonction de l'intensité de fluorescence DsRed. Deux conditions, les cellules transfectées avec le siCD9 contrôle en présence de doxycycline (courbe jaune) et avec le siDsRedB en présence de doxycycline (courbe rouge).

## 2.2.2 - Etude de l'insertion d'une séquence PolyA dans pStan

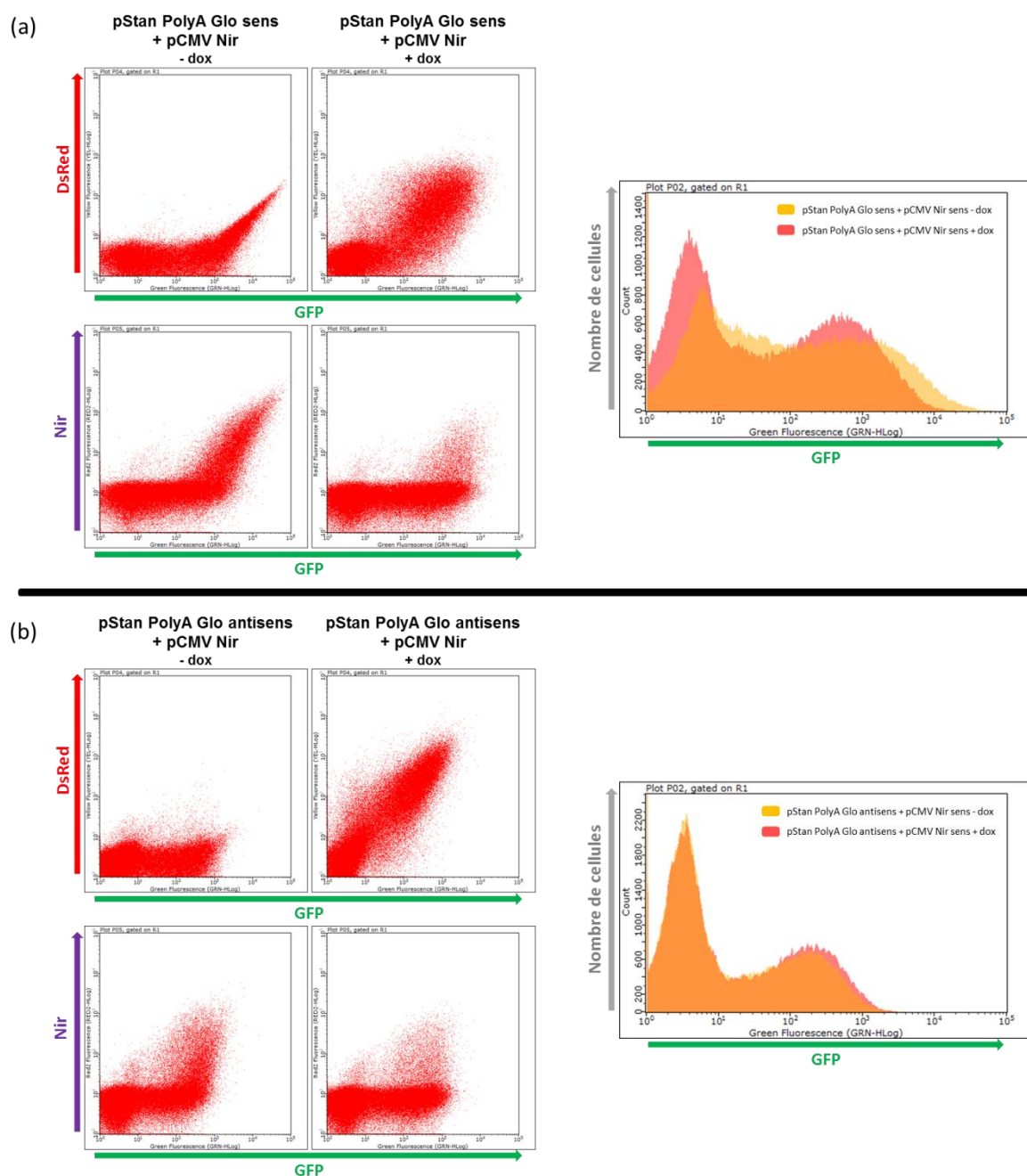
Les données précédentes font écho à d'autres résultats obtenus *en transitoire* avec des constructions pStan où une séquence de polyadénylation entre les deux séquences codantes a été insérée. Deux types de constructions ont été créés avec soit la séquence de polyadénylation en orientation sens soit en orientation antisens vis-à-vis du gène GFP.



Les résultats obtenus présentés dans la figure 40a montrent que lorsque la séquence est orientée en sens, l'expression de la GFP est bien meilleure. Cela vient sans doute du fait que la polyadénylation à ce niveau donne un transcrit plus court et sans doute plus stable que lorsque la polyadénylation se fait au-delà du promoteur pBi. D'autre part, avec ce site de polyadénylation entre les deux régions codantes, on peut constater que l'expression de la construction contrôle pCMV Nir sens est bien meilleure qu'avec une construction pStan classique. Cela est clairement un signe de la perte d'une certaine complexité dans les transcrits issus du promoteur CMV qui permettait de mettre en place un silencing non spécifique. Néanmoins avec cette construction, l'ajout de la doxycycline au milieu de culture, comme c'est le cas avec le pStan classique, déclenche un silencing sur l'expression de la GFP mais également sur la construction contrôle. Ainsi l'induction de l'activité du promoteur pBi permet de réinstaller en partie un silencing non spécifique.

Les résultats les plus intéressants viennent de la construction pStan où la séquence de polyadénylation se trouve cette fois-ci en antisens par rapport au gène GFP (Fig. 40b). En absence de doxycycline, le profil est peu différent d'une construction pStan classique mise à part une expression sensiblement plus forte du contrôle pCMV Nir sens, signe que le silencing non spécifique n'est pas uniquement dû aux produits du promoteur CMV mais peut être également dû à un taux de fuite du promoteur pBi. En présence de doxycycline on observe, comme c'était également le cas aussi pour la GFP, une meilleure expression de la DsRed avec un transcrit DsRed plus court. Cela laisse penser que, dans le pStan classique où le transcrit DsRed est plus long et comporte la partie GFP antisens, ce dernier est peut-être moins apte à la traduction. Là où les données deviennent intéressantes, c'est qu'en présence de cette séquence de polyadénylation en antisens qui empêche la transcription au-delà et donc la transcription de l'antisens GFP, nous n'observons pas de silencing. Ainsi l'absence de transcription au-delà de cette séquence semble supprimer le silencing présent dans les autres constructions de ce type. Néanmoins, cette donnée ne nous permet pas de conclure sur la nature du silencing observé.





**Figure 40 : Etude des constructions pStan PolyA Glo sens et antisens en transfection transitoire**

Expériences réalisées dans les cellules HEK 293 3G Tet-on, transfection de 0,25 µg de plasmide et analyse 48h après transfection en cytométrie en flux (GUAVA). La doxycycline est ajouté le jour de la transfection (1 µg/mL). (a) Profils de fluorescence et histogrammes représentant la fluorescence GFP de la construction pStan PolyA Glo sens en absence et en présence de doxycycline dans le milieu de culture. (b) Profils de fluorescence et histogrammes représentant le nombre de cellules en fonction de l'intensité de fluorescence GFP de la construction pStan PolyA Glo antisens en absence et en présence de doxycycline dans le milieu de culture.

En effet, deux hypothèses sont envisageables : un effet de l'antisens lui-même ou simplement de sa transcription. Il est cependant clair que ce silencing en *cis* n'a pas pour origine une régulation de type post-transcriptionnel.

L'intégration de la construction pStan PolyA Glo antisens a été réalisée pour permettre une

analyse en l'absence de silencing non spécifique. Ainsi comme pour la construction pStan, trois clones ont été réalisés (Fig. 41). A la différence de ce qui a été observé avec la construction de départ, ici deux clones répondent à l'induction de l'antisens par une augmentation de l'expression. Ceci est compatible avec une perte du silencing qui laisserait comme seule réponse l'effet de transactivation à partir du promoteur pBi. Cependant, dans un autre clone on observe toujours un silencing ce qui ne permet pas de conclure définitivement sur l'effet de la séquence de polyadénylation. Un point important est que la séquence de polyadénylation n'est pas un site de terminaison et que celle-ci se fera en fonction de la capacité d'élongation de la polymérase. Aussi, en plus de l'organisation des constructions intégrées dans chacun des sites, l'environnement chromatinien pourrait aussi jouer un rôle sur l'efficacité de la terminaison de la transcription.

Clone	Effet observé sur la GFP de l'induction de l'antisens	Expression de la DsRed après induction
HEK 293 pStan PolyA Glo antisens clone 8	-55%	Forte
HEK 293 pStan PolyA Glo antisens clone 9	+60%	Moyenne
HEK 293 pStan PolyA Glo antisens clone 21	+87%	Forte

**Figure 41 : Tableau de l'effet des clones HEK 293 pStan PolyA Glo antisens**

Le pourcentage de l'effet sur la GFP est obtenu par comparaison de l'intensité de la fluorescence après 48h de culture de ces cellules entre les conditions avec et sans doxycycline supplémentée au milieu de culture.

### 2.2.3 – Ajout de l'élément PRC2 à la construction pStan intégrée

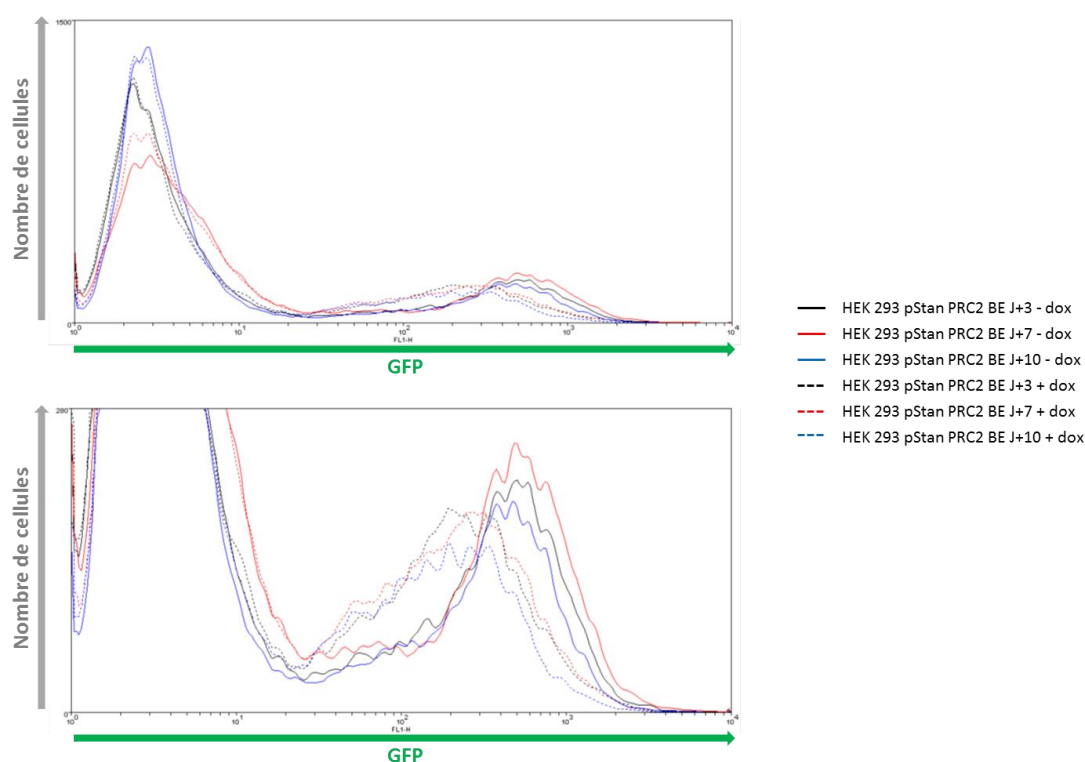
Le silencing ne se maintenant pas au-delà de l'expression de l'antisens, il a été essayé de réaliser un silencing plus stable en insérant le « Binding Element PRC2 » au sein de la construction pStan. Il est en effet plus probable, dans ce contexte en *cis*, que le transcrit antisens se retrouve à proximité de sa cible et puisse amener les protéines effectrices pour la mise en place d'un silencing transcriptionnel. Ainsi, l'élément permettant le recrutement du complexe Polycomb par l'ARN est inséré dans la construction pStan entre les deux séquences codantes dans une orientation antisens vis-à-vis du gène de la GFP. La construction pStan PRC2 BE alors créée est ensuite intégrée au génome des cellules HEK 293 Tet-on. Ici, contrairement aux cas précédents, les clones ayant intégré la construction ne sont pas isolés mais regroupés pour former un « pool » afin de pouvoir tester la réponse de nombreux clones. Ainsi, le profil d'expression de la lignée est très hétérogène et ressemble à





# Résultats

l'expression des constructions en transitoire (Fig. 40). Les résultats montrent que malgré la présence de l'élément pouvant permettre le recrutement du complexe Polycomb, le silencing permis par l'ajout de la doxycycline et donc l'activation du promoteur pBi n'est pas différent de celui d'une construction pStan classique. En particulier, le maintien de la doxycycline dans le milieu de culture ne modifie pas le silencing. Comme pour le pStan classique le maximum du silencing observé est atteint dès le premier point (3 jours). On observe cependant quelques différences entre les résultats. Les petites différences d'expression aux différents temps traduisent les variations de l'expression en fonction de la densité cellulaire et sont observables en présence et en l'absence de doxycycline.



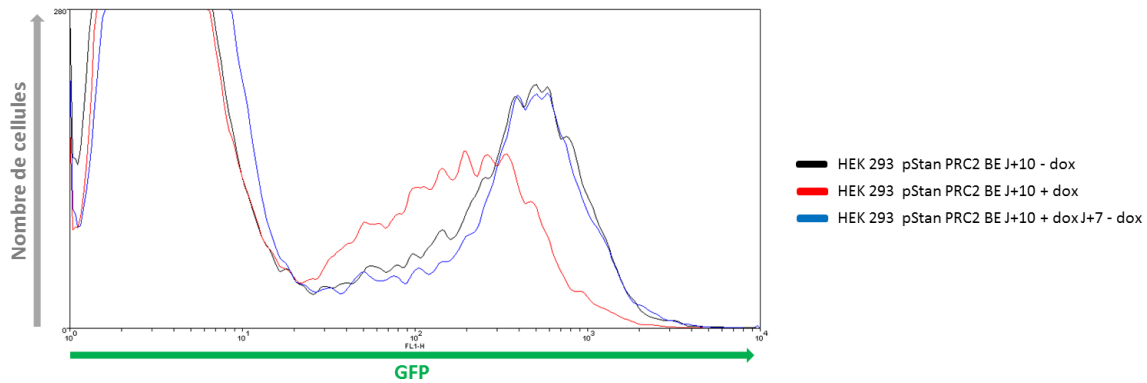
**Figure 42 : Cinétique longue des cellules HEK 293 Tet-on pStan PRC2 BE**

Expérience réalisée dans les cellules HEK 293 Tet-on pStan PRC2 BE. Les résultats sont obtenus sous forme d'histogrammes de fluorescence GFP après analyse en cytométrie en flux (FACS Calibur) des cellules cultivées pendant 3, 7 ou 11 jours en milieu de culture supplémenté ou non par de la doxycycline. Les deux histogrammes représentent les mêmes données mais à des échelles différentes pour une meilleure visibilité.

Afin de vérifier si l'élément permettant le recrutement du complexe Polycomb a un effet sur l'organisation de la chromatine et induit une mémoire épigénétique, la doxycycline est retirée du milieu de culture à J+11 et les cellules sont cultivées sans doxycycline pendant une



semaine. Comme on le constate sur la figure 43, le silencing disparaît en l'absence de doxycycline et l'expression de la GFP retrouve son profil initial. Ce résultat s'apparente à celui obtenu avec la construction pStan de départ pour laquelle aucune mémoire du silencing n'a été observée (Fig. 38).



**Figure 43 : Cinétique longue des cellules HEK 293 Tet-on pStan PRC2 BE**

Expérience réalisée dans les cellules HEK 293 Tet-on pStan PRC2 BE. Les résultats sont obtenus sous forme d'histogrammes de fluorescence GFP après analyse en cytométrie en flux (FACS Calibur). Les cellules sont cultivées pendant 10 jours dans un milieu avec ou sans doxycycline. Dans le cas de la courbe bleue, les cellules sont cultivées pendant 10 jours dans un milieu avec doxycycline puis les cellules sont cultivées 7 jours de plus dans un milieu sans doxycycline.

## Discussion

Le projet de thèse reposait sur trois éléments : i) l'utilisation de constructions indicatrices a priori dépourvues de motifs spécifiques de régulation par les ARN antisens, ii) une analyse dans un premier temps par des transfections transitoires, iii) une analyse de l'expression au niveau de chaque cellule au moyen de la cytométrie. Une telle approche était destinée à permettre l'analyse de la contribution des différents mécanismes possibles en partant des contextes les plus simples (comme une expression épisomique avec une chromatinisation partielle) et en introduisant progressivement d'autres éléments de régulation. De plus la cytométrie permet de distinguer entre une réponse dominante et l'existence de réponses spécifiques dans des sous-populations. Cependant, ce programme s'est trouvé confronté à plusieurs niveaux de difficultés qui en ont réduit la mise en œuvre.

Les premières difficultés rencontrées ont été celles liées à plusieurs niveaux de réponses aspécifiques lors de l'expression transitoire.

### 1 – La réponse en transitoire et les effets non spécifiques

Dans le schéma de principe de notre approche, nous utilisons au minimum deux protéines fluorescentes afin de pouvoir suivre l'expression du gène cible et l'expression de l'ARN régulateur. Une troisième protéine peut servir de contrôle interne pour suivre l'efficacité de transfection et normaliser les signaux. Du fait du format des expériences qui sont réalisées en plaque douze puits, le plus souvent à quatre points par condition, et de la robustesse du modèle cellulaire - les cellules HEK 293 étant particulièrement transfectables - les variations d'expression en transitoire ne dépassent pas 20 % entre les différentes conditions expérimentales. Ainsi le contrôle de transfection peut-il être en partie utilisé pour suivre des effets de silencing non spécifique du moins si ceux-ci sont suffisamment marqués. De fait, la présence de constructions permettant l'expression d'un ARN sens et antisens déclenche systématiquement un silencing affectant les constructions co-transfectées. Une difficulté potentielle de notre approche est que toutes les protéines fluorescentes que nous avons utilisées sont apparentées entre elles et donc présentent une homologie de séquence (47% d'homologie entre GFP et DsRed). On peut toutefois remarquer que l'expression de l'ARN



antisens de l'une de ces protéines n'a pas d'effet sur l'expression des autres tant que l'ARN sens correspondant n'est pas exprimé. Sans pouvoir complètement exclure que les homologies de séquence puissent jouer un rôle dans les silencings que nous avons observés, ce résultat établit que seul un duplex parfait induit cette réponse.

### ***La réponse aux ARN double brin***

La capacité des ARN double brin à induire des réponses cellulaires, ici désignées collectivement par réponses de type interféron, a été mise en évidence bien avant la découverte de l'interférence par l'ARN. Ce phénomène a d'ailleurs été initialement interprété comme rendant impossible l'utilisation des ARN double brin pour induire une interférence par l'ARN en dehors des cas où ces réponses sont absentes (comme les cellules souches embryonnaires). Cependant, l'expression d'un ARN antisens a été utilisée pour inhiber l'expression du gène correspondant et ainsi permettre des études fonctionnelles, suggérant que l'induction des réponses de type interféron n'est pas systématique ou dépend du mode de production des ARN double brin (expression d'un transcrit sens et antisens vs introduction d'ARN double brin). Plus récemment, le groupe de Petr Svoboda a montré que l'expression d'un ARN en tige-boucle n'induisait pas de réponse majeure dans des souris transgéniques bien que le produit du transgène comprenne une région parfaitement appariée de 520 paires de bases (Nejepinska *et al*, 2012). Plus spécifiquement, les souris sont viables, la voie interféron n'est pas activée et des petits ARN interférants ne sont détectés que dans les oocytes. Ainsi l'expression d'un ARN double brin (dans ce cas-ci intramoléculaire) n'entraîne pas nécessairement de réponse cellulaire forte.

### ***L'activation de la kinase PKR***

Plusieurs voies de réponse aux ARN double brin ont ainsi été caractérisées. La publication du groupe de Petr Svoboda qui est intervenue pendant que je réalisais les expériences de transfection transitoire a focalisé notre attention sur l'une de ces voies, celle de la kinase activée par les ARN double brin, PKR (Nejepinska *et al*, 2014). Cette étude, en prolongement de celle sur les souris transgéniques, visait à caractériser plus complètement la réponse des



cellules à l'expression d'un ARN en tige-boucle de grande taille à partir d'une construction plasmidique. L'expression de cet ARN déclenche une réponse aspécifique qui affecte des constructions indicatrices luciférase co-transfectées et est donc *a priori* similaire à la nôtre. Ils ont pu de plus montré que la kinase PKR était l'acteur majeur de la réponse aspécifique en utilisant soit un inhibiteur chimique soit un petit ARN interférant. Toutefois, en dépit des similitudes entre nos deux modèles, nous n'avons pas observé une implication forte de PKR dans le silencing aspécifique dans les cellules HEK 293, suggérant l'implication d'autres réponses. La différence majeure entre les deux études vient de la genèse des ARN double brin, soit intramoléculaire dans leur cas, soit bimoléculaire dans le nôtre. On peut donc envisager que la différence de réponse entre les deux modèles reflète ce mécanisme de formation de l'ARN double brin avec des concentrations *a priori* plus élevées dans le cas intramoléculaire. Mais ceci ne peut tout expliquer puisque dans notre étude nous pouvons étudier la réponse des cellules en fonction du niveau d'expression de la cible et du régulateur. Une autre différence qui pourrait expliquer l'activation plus forte de PKR par la tige boucle est que le transcrit qui la porte est en principe associé à la machinerie de traduction puisqu'il s'agit d'un ARN messenger EGFP. Ainsi la phosphorylation d'eIF2a pourrait-elle être faite sur l'ARN en situation de traduction. A l'inverse, dans notre cas, la région double brin s'étend jusqu'à l'AUG et on peut supposer que le duplex n'est plus traductible.

PKR a été décrite comme pouvant agir en *cis* et en *trans* sans que l'importance relative de ces deux phénomènes ait été complètement déterminée. D'une part, à travers la phosphorylation du facteur de traduction eIF2a, il existe la possibilité d'une action globale sur l'efficacité de traduction des messagers cellulaires (Garcia *et al*, 2006). Sans exclure que certains messagers soient plus sensibles à l'activité de PKR ceci peut conduire à un silencing global en *trans*. D'autre part, comme cela a été décrit dans le cas de l'interféron gamma, l'activation de la kinase PKR pourrait agir principalement en *cis* sur la traduction des ARN qui contiennent une tige-boucle (Ben-Asouli *et al*, 2002). On peut remarquer que l'étude de Nejepinska comprend un volet de cytométrie, l'ARN en tige boucle étant intégré en aval de la région codante de l'EGFP. Les cellules ayant une fluorescence verte significative, la présence en *cis* d'une tige boucle ne suffit pas à rendre impossible la traduction. Cependant, en comparaison avec l'expression de la construction sans tige boucle on observe un silencing



important de l'expression de l'EGFP. De plus, le profil de fluorescence évoque fortement ceux que nous observons lors d'un silencing induit par l'ARN antisens avec une distribution bimodale. Ainsi, bien que l'activation de la kinase PKR n'explique pas les réponses observées dans notre modèle, il semble que dans les deux cas il y ait une réponse distincte entre les forts expresseurs et les moyens expresseurs. Ceci peut refléter le fait que dans les deux modèles il y a une corrélation entre l'expression de la cible et de l'ARN régulateur ; directe dans le cas de Nejepinska où les deux sont portés par le même transcrit, globale dans notre cas où les deux promoteurs sont soit identiques soit d'expression corrélée dans le cas CMV et pBi.

### ***Autres voies de réponse aux ARN double brin***

Au-delà de PKR, d'autres voies de réponse aux ARN double brin pourraient être activées dans notre cas. C'est dans cette idée qu'a été testée l'inhibition de la RNaseL, une autre branche de la réponse interféron qui permet la dégradation des ARN double brin préalablement marqués par la 2'-5' Oligoadenylate Synthase. Mais, l'inhibition de l'expression de la RNaseL n'a pas d'effet notable sur le silencing aspécifique. D'autres protéines, impliquées dans l'immunité innée, peuvent détecter la présence d'ARN double brin dans le cytoplasme chez les mammifères. C'est le cas des récepteurs de la famille RLR (RIG-I like Receptors) (Yoneyama *et al*, 2015). RIG-I interagit avec l'extrémité 5' d'un ARN double brin, la présence d'un phosphate en position  $\beta$  étant un élément clé de l'interaction. Aussi, les petits ARN double brin produits par Dicer qui ne disposent que d'un seul phosphate en 5' n'activent-ils pas RIG-I. MDA5 interagit avec de longs ARN double brin et pas avec leur extrémité 5' du fait d'une différence dans le domaine C-terminal qui chez RIG-I coiffe le transcrit situé dans le domaine hélicase. Globalement, il est estimé que les ARN double brin de grande taille (> 500 pb ou 2 000 pb suivant les études (Kato *et al*, 2008)) activent MDA5, tandis que ceux qui sont plus petits activent RIG-I.

Enfin, la présence d'ARN double brin pourrait induire une réponse de type interférence par l'ARN, qu'elle soit cytoplasmique ou nucléaire. Bien qu'il existe peu de données en faveur d'une telle réponse dans des cellules somatiques, nous avons testé l'implication d'Ago2 mais



sans observer d'effet. Les homologies de séquence entre les protéines fluorescences utilisées sont d'ailleurs trop faibles pour qu'un mécanisme spécifique comme l'interférence puisse expliquer le silencing observé sur les contrôles. Au total, il apparaît que dans les expériences en transitoire l'expression d'ARN sens et antisens conduit à des effets non spécifiques importants qui masquent une réponse spécifique du gène ciblé, si elle existe.

### ***La différence de réponse entre transcrits génomiques et épisomiques***

Une piste sur la nature réelle de cette réponse aspécifique réside dans les observations réalisées en clones stables. En effet, il existe un contraste fort entre le cas transitoire et intégré qui suggère une nature antivirale de la réponse mise en place. Lorsque les constructions indicatrices sens et antisens sont intégrées, l'expression de l'antisens ne déclenche plus aucune réponse aspécifique. Ce type d'observation a également été faite pour les mécanismes PKR-dépendant ou RNaseL-dépendant, où ces derniers vont être capables d'affecter l'expression de vecteurs viraux ou non viraux (plasmides) mais n'auront pas d'effet sur l'expression d'un gène endogène (Terenzi *et al*, 1999). Cela nous informe sans doute sur la nature de ce silencing mais soulève plusieurs questions comme la façon dont ces mécanismes sont capables de différencier un ARN épisomique d'un ARN génomique. Il faut qu'il existe une différence notable entre ces deux types de transcrits qui pourrait être au niveau des protéines associées, qui sont peut être de nature différentes de façon à ce que la cellule soit capable de différencier le soi du non-soi, ou des propriétés intrinsèques de l'ARN, par exemple la taille de la queue PolyA. Pour aborder ce dernier cas il serait intéressant d'analyser la longueur des queues PolyA des deux classes de transcrits.

Une autre donnée intéressante est que ce mécanisme peut être déclenché par un duplex d'ARN formé par un sens intégré et un antisens exprimé en transitoire sans que cela affecte l'expression du gène sens. Cela confirme que l'ARN messager sens peut être le déclencheur de la réponse aspécifique sans faire partie des cibles, alors que si le même ARN était exprimé depuis un épisode il serait soumis au silencing. L'expérience inverse a été réalisée, où l'ARN antisens est exprimé en intégré avec une cible, elle exprimée en transitoire. Dans ce cas le silencing aspécifique est plus modéré, voire absent, pour certains clones stables exprimant l'antisens. Cette observation laisse à penser qu'il y a peut-être dans certains contextes une



dissymétrie et qu'il est possible que ce soit l'ARN antisens qui dicte la mise en place d'une réponse aspécifique. Dans l'idée que le mécanisme impliqué est une réponse antivirale, il est envisageable que l'ARN antisens issu du génome ciblant un transcrit sens issu de l'épisome soit reconnu comme une régulation légitime de la cellule alors que le cas inverse signerait l'influence d'une expression extérieure et déclencherait la réponse non spécifique. Une étude à plus grande échelle serait nécessaire pour déterminer s'il existe vraiment une différence dans la capacité des transcrits sens et antisens à induire une réponse.

### ***Les transcrits internes aux plasmides***

Une autre difficulté a été rencontrée dans l'utilisation du modèle transitoire en plus de la réponse aspécifique. La co-transfection de plusieurs plasmides induit des « interférences » entre ceux-ci. Cette observation très courante est souvent attribuée à la compétition entre les promoteurs pour les facteurs de transcription ou de manière plus globale pour la machinerie de transcription. Cependant, un autre mécanisme a été proposé par l'équipe de P. Svoboda et repose sur la transcription « illégitime » des plasmides non intégrés. L'étude du transcriptome des plasmides en transfection transitoire effectuée a montré que par exemple dans les constructions commerciales de type pCMV, la cassette de résistance à la kanamycine générerait des ARN sens et antisens pouvant déclencher une réponse cellulaire (Nejepinska *et al*, 2012). Ces transcrits sens et antisens seraient capables de réprimer de façon basale, avant même l'ajout de l'antisens dans notre modèle, l'expression des constructions plasmidiques. Ils pourraient aussi participer aux interférences entre plasmides. Les données présentées par l'équipe de Petr Svoboda montrent que ce phénomène est PKR-dépendant. Cette donnée est confirmée par nos expériences d'inhibition de PKR, que ce soit par l'inhibiteur C16 ou le siARN, où l'on observe une augmentation de l'expression basale des constructions utilisées. Ainsi, bien que PKR ne semble pas jouer un rôle déterminant dans le silencing non spécifique que nous observons en réponse à l'expression d'un ARN antisens, elle est néanmoins activée dans nos expériences de transfection transitoire. L'un des objectifs de ces expériences en transitoire était d'évaluer l'importance d'un conflit transcriptionnel lorsque les unités de transcription sens et antisens se recouvrent. La construction pStan permet en principe d'aborder cette question. Cependant le silencing





observé en réponse à l'induction de la transcription de l'antisens est du même ordre que dans les autres expériences y compris lorsque l'antisens est exprimé depuis un autre plasmide. De plus, il se manifeste aussi sur les constructions contrôles confirmant qu'il s'agit au moins pour une grande part d'un effet non spécifique. L'importance de ces effets non spécifiques ne permet pas d'évaluer précisément l'impact d'un conflit transcriptionnel. On peut cependant remarquer (Figure 22) que les cellules qui expriment le plus fortement l'ARN antisens (détecté via la DsRed) sont aussi celles qui expriment le plus l'ARN sens (détecté via l'expression de la GFP). Ceci suggère que même dans des cellules qui expriment des niveaux très élevés d'ARN sens et antisens le conflit n'est pas suffisamment important pour induire une expression mutuellement exclusive. Cette conclusion repose sur la capacité d'analyser l'expression au niveau de chaque cellule mais doit être tempérée par le fait que la transfection introduit de nombreuses copies du plasmide. On pourrait donc envisager que la forte expression des deux protéines provienne en fait de la transcription de plasmides différents. Cependant, les expériences réalisées avec une construction pStan linéarisée dans l'ossature du plasmide ne montrent aucune diminution de la GFP en réponse à l'induction de la transcription de DsRed/GFPas. Aussi en dépit des nombreuses limitations des expériences en transitoire, l'ensemble des données est-il en faveur d'un rôle très limité des conflits transcriptionnels même en situation de forte expression.

## 2 – L'étude de l'effet de l'ARN antisens en clones stables

### ***Cas d'un ARN antisens exprimé en trans***

Comme attendu les effets non spécifiques sont nettement diminués et deviennent inférieurs à 20 % lorsque la construction indicatrice est intégrée dans le génome (Figure 26), ce qui rend possible d'analyser un silencing spécifique (sur la cible complémentaire de l'ARN antisens). Cependant, que l'ARN antisens soit transcrit à partir d'un plasmide en transitoire ou d'une construction intégrée, aucun effet significatif n'a été observé sur la cible directe. Cette absence d'effet en *trans* n'est pas en soi une surprise complète puisque même chez la levure *S. pombe* où les régulations transcriptionnelles induites par des ARN double brin ont été étudiées en détail, la possibilité d'agir en *trans* semble réduite. Ce d'autant plus que dans notre démarche, nous n'introduisons pas un ARN double brin mais un ARN antisens qui



doit ensuite s'hybrider avec la transcrit complémentaire afin d'induire une réponse. De plus, si l'on s'intéresse à une réponse nucléaire, il faudrait probablement que l'hybridation ait lieu avec le transcrit naissant de la cible. Par contre on aurait pu s'attendre à une réponse post-transcriptionnelle si le duplex se forme dans le cytoplasme et a minima rend l'ARNm non traductible.

Plusieurs éléments de réponse peuvent être envisagés pour expliquer l'absence d'effet observée. Tout d'abord, par rapport aux transcrits cellulaires, nos constructions sont particulièrement simples : de petite taille et dépourvues d'intron. Ainsi, le gène de l'EGFP n'est constitué que de 700 paires de bases et est donc rapidement transcrit (~ 30 s sur la base d'une vitesse d'élongation de 1500 nt/min (Conaway *et al*, 2000)). De plus, la présence de signaux de polyadénylation particulièrement efficaces conduit probablement à une maturation rapide des transcrits, suivie de leur dispersion dans le nucléoplasme et de leur transport vers le cytoplasme. De la même façon, l'ARN antisens est lui aussi court quand on le compare aux antisens naturels caractérisés chez les mammifères dont l'unité de transcription s'étend le plus souvent sur plusieurs kilobases (Yu *et al*, 2008 ; Rinn *et al*, 2007 ; Brown *et al*, 1991). Ainsi, le temps de rétention dans le noyau est sans doute faible pour les deux transcrits ce qui réduit la probabilité d'une rencontre intranucléaire. Plusieurs approches ont été testées afin d'augmenter leur résidence dans le noyau avec l'utilisation d'une version de la GFP fusionnée avec la  $\beta$ -globine (taille du transcrit primaire 2 kb et nécessitant deux épissages) ou en supprimant la séquence de polyadénylation dans la construction antisens ce qui pourrait conduire à un transcrit primaire beaucoup plus long et à une absence d'exportation en l'absence d'une queue PolyA (Shinagawa *et al*, 2003 ; Gullerova *et al*, 2012). Néanmoins, ces modifications n'ont pas permis la mise en place d'un silencing spécifique lié à l'antisens. Dans le même objectif de permettre au transcrit de rester le plus longtemps possible dans le noyau, une partie de notre étude s'est intéressée à la protéine PAF1 ou « RNA-Polymerase-Associated Factor 1 ». Cette dernière participe à l'élongation lors de la transcription. Il a été montré chez *S. pombe* que l'inhibition de cette protéine permettait la mise en place d'un silencing épigénétique médié par des petits ARN (Kowalik *et al*, 2015). Il semblerait que chez *S. pombe*, la présence du complexe PAF1 empêche le ciblage du transcrit naissant par ces mêmes petits ARNs. Dans notre modèle, l'inhibition de PAF1 a été testée pour vérifier si le ralentissement de l'élongation augmentait



l'accessibilité du transcrit naissant et permettait la mise en place d'un silencing. Cependant, le seul effet visible a été une diminution de l'expression, probablement du fait d'une élongation de la transcription moins efficace sans conduire à un silencing spécifique.

Sans exclure que le temps de résidence du transcrit naissant au voisinage du gène soit un paramètre important, nos résultats suggèrent que d'autres aspects interviennent dans la réponse en *trans*. Le plus évident est celui de l'organisation spatiale du noyau, le transgène et la construction antisens étant probablement transcrits dans des usines de transcription différentes (Osborne *et al*, 2004). Cette hypothèse a été validée au moins pour certains ARN non codants agissant en *trans*. C'est le cas par exemple de l'ARN HOTAIR qui a pour cible des gènes situés sur un autre chromosome que celui sur lequel il est transcrit, mais qui en réalité se trouvent à proximité dans l'organisation spatiale du noyau (Denholtz *et al*, 2013 ; Noordermeer *et al*, 2014). Il se peut donc que la proximité des deux unités de transcription dans le noyau soit un facteur primordial pour la mise en place d'un silencing ; ceci reflète probablement l'importance de la concentration locale de l'ARN antisens au niveau de son gène cible. Dans le cas du long ARN non codant roX, on observe ainsi que bien qu'il puisse se fixer sur l'ADN au niveau des éléments CLAMP, seuls ceux qui sont proches spatialement sont occupés (Soruco *et al*, 2013). L'importance de l'organisation spatiale et de la concentration locale a aussi été discutée dans le cas du silencing transcriptionnel médié par les petits ARN double brin chez *S. pombe*. Dans ce mécanisme les petits ARN nouvellement générés par la fixation du complexe effecteur RITS à un transcrit naissant vont pouvoir agir en *cis* mais pas en *trans* sur les cibles partageant cette séquence (Buhler *et al*, 2006).

L'ensemble de ces résultats suggère que l'action en *trans* d'un ARN antisens nécessite des conditions particulières. De fait, la très grande majorité des ARN antisens agit dans un contexte en *cis* et seule l'action en *trans* de PTENP1-AS a été bien documentée (Johnsson *et al*, 2013). Dans ce cas les deux gènes ne sont pas situés sur le même chromosome (PTEN en 10q23, PTEN-AS en 9p13) mais leur distance dans le noyau n'a pas été caractérisée.

Plus inattendue est l'absence d'effets post-transcriptionnels sur la cible spécifique. En effet, au-delà de la formation d'ARN double brin dans le noyau, ceux-ci pourraient se former dans le cytoplasme et induire plusieurs réponses sur la cible directe : interférence par l'ARN ou inhibition de la traduction (« effet antisens »). Pourtant, aucun silencing significatif n'est observable sur la cible spécifique, le léger silencing (< 20 %) observé dans la Figure 27 affecte



aussi bien la DsRed que la GFP. Pourtant, des effets non spécifiques sont clairement visibles sur une construction contrôle exprimée en transitoire, supprimant pratiquement complètement son expression.

Notre modèle de constructions indicatrices qui était élaboré pour une étude en transitoire de l'antisens s'est montré être un peu trop simple pour une analyse en clone stable. Sa simplicité et le manque de contrôle sur les sites d'intégration ne nous permettent pas une analyse complète. Il est envisageable d'approfondir la question en passant sur l'étude d'un gène cellulaire et de cibler les sites d'insertion de notre construction via l'utilisation de CRISPR/Cas9.

### ***Cas d'un ARN antisens exprimé en cis***

Les cis-NAT correspondent à l'organisation la plus fréquemment observée chez les mammifères dans laquelle le transcrit antisens est issu du même fragment d'ADN que le transcrit sens. Plusieurs types d'organisation sont néanmoins possibles selon le niveau de recouvrement entre les deux unités de transcription et l'organisation en exon et intron de celles-ci. Pour étudier l'action des ARN antisens en *cis* j'ai réalisé une construction qui conduit à un recouvrement maximal entre les unités de transcription puisque pour chaque transcrit le site de polyadénylation est placé en aval du promoteur convergent. Ainsi, les possibilités d'action de la transcription antisens mais aussi des transcrits antisens sont-elles maximales. Par contre, dans leur première forme, ces constructions sont potentiellement dépourvues de motif de recrutement de complexes modificateurs de la chromatine. Enfin, l'utilisation d'un promoteur régulé par la doxycycline permet d'aborder la dynamique de la réponse.

L'analyse de plusieurs clones met en évidence deux types de réponse à l'induction de la transcription antisens ; dans certains clones on observe un silencing du gène cible tandis que dans d'autres on observe une induction. Cette dualité de réponse est à l'image de l'absence de relation entre l'expression des transcrits sens et antisens au niveau du génome. Comme évoqué dans la problématique, ceci reflète très probablement les effets opposés d'un silencing et d'une transactivation liée à l'activation du promoteur antisens. Nos résultats suggèrent cependant que, quels que soient les mécanismes impliqués, ceux-ci ne sont pas



clairement dominants puisque la réponse dépend du contexte. Une difficulté importante de l'analyse qui a été faite ici est l'utilisation de la transfection pour obtenir des clones stables. Ceci introduit deux sources de variabilité entre clones, le site d'intégration (le plus souvent celle-ci a lieu dans un seul site) et le nombre et l'organisation des copies intégrées. Le développement de la technologie CRISPR/cas9 dans les cellules de mammifère permet dorénavant de remplacer ce mode d'intégration au hasard par une intégration par recombinaison homologue dans un site prédéfini du génome. Cette approche ne pouvait jusqu'à présent être mise en œuvre que dans les cellules souches embryonnaires où l'efficacité de la recombinaison homologue est la plus importante, c'est par exemple ce qui a été réalisé pour l'étude de la régulation de Igf2r par Air (Latos *et al*, 2012). Il est à noter que dans cette étude l'introduction d'un site de polyadénylation est considérée comme équivalente à celle d'un site de terminaison de la transcription. L'induction d'une coupure de l'ADN par CRISPR/cas9 permet maintenant de mettre en œuvre cette stratégie dans des cellules somatiques et donc par exemple dans les cellules HEK 293 Tet-on que nous avons utilisées. Ceci permet de n'intégrer qu'une seule copie de la construction et, de plus, dans une configuration prédéfinie, par contre il restera à choisir le site d'intégration.

Dans les clones où un silencing est observé, celui-ci n'affecte pas l'expression d'une construction contrôle exprimée en transitoire. Ainsi, bien que les ARN sens et antisens soient transportés dans le cytoplasme puisque l'expression de la GFP et de la DsRed est observée dans les cellules, ceci n'induit pas les effets non spécifiques décrits plus haut. On peut noter que dans les clones stables les niveaux d'expression sont globalement plus faibles que ceux des forts expresseurs obtenus en transitoire. Cependant, même si les analyses tendent à privilégier les forts expresseurs, il est apparent que les effets non spécifiques concernent aussi des niveaux d'expression intermédiaires qui correspondent à ceux observés en clones stables. Compte tenu que lorsqu'une construction sens et une construction antisens sont intégrées dans le génome on n'observe pas non plus d'effets non spécifiques, l'interprétation la plus simple de ces résultats est que l'origine des transcrits (épissomique ou génomique) détermine la mise en place de ces effets. Il a déjà été évoqué que l'on pouvait observer des effets non spécifiques dans le cas d'une expression mixte (un transcrit génomique et un épissomique) mais que l'amplitude des effets était variable. Ainsi, l'origine du transcrit conditionne non seulement sa sensibilité à ces effets non spécifiques mais aussi



leur mise en place.

Dans les clones où l'on observe un silencing en réponse à l'expression d'un ARN antisens le niveau d'expression de ce dernier ne semble pas jouer un rôle déterminant au premier abord. En effet, les cellules qui expriment le niveau le plus fort de DsRed sont aussi celles qui expriment le niveau le plus fort de GFP. L'interprétation de ce résultat doit cependant être tempérée par la possibilité que le silencing de la GFP soit contrebalancé par un effet de transactivation. On peut néanmoins remarquer que les faibles expresseurs de la DsRed sont enrichis en forts expresseurs de la GFP, en accord avec un effet dose dans la mise en place du silencing. Une induction prolongée ne change que peu le profil observé permettant cependant une expression plus forte de la DsRed. Lorsque l'on prive les cellules de doxycycline après les y avoir exposées, le niveau d'expression de la GFP revient à celui d'origine en fonction de la diminution de l'expression de l'antisens. La doxycycline s'accumulant dans les membranes de la cellule, la purge peut prendre plusieurs jours en fonction de la dose utilisée. Ainsi le silencing observé n'induit-il pas d'effet mémoire. Comme par ailleurs le silencing d'Ago2 n'influe pas sur la réponse observée, l'interprétation la plus simple est que le silencing induit par la transcription antisens est dû à un conflit transcriptionnel. Il reste alors à comprendre pourquoi le conflit transcriptionnel n'est pas observé dans les expériences en transitoire en dépit de niveaux d'expression plus élevés. Une possibilité est que la chromatinisation incomplète des constructions indicatrices en transitoire soit précisément la cause de cette différence. En accord avec l'abondance des transcrits « illégitimes » observés dans les expériences transitoires, on peut supposer que la transcription est plus facile à mettre en place sur une matrice ADN faiblement associée à des protéines. Dans ce contexte, la fixation de facteurs de transcription pourrait rapidement déboucher sur le recrutement du complexe de transcription. De plus, la transcription elle-même pourrait être plus rapide. Or si le concept de conflit transcriptionnel est facile à définir il est beaucoup plus difficile de lui donner une base moléculaire. En effet même les gènes les plus exprimés – si l'on exclut la transcription par l'ARN Polymérase I des gènes ribosomiques – ne sont transcrits qu'épisodiquement. A titre d'ordre de grandeur, un ARN messager très abondant peut s'accumuler jusqu'à quelques centaines de molécules par cellule. Si sa demi-vie est de 12 h, le gène correspondant sera transcrit toutes les cinq minutes (pour un niveau d'accumulation de 300 molécules). Compte tenu de la taille de la construction pStan il en



résulte que la transcription sens n'occupera le gène que pendant environ 20 % du temps. Un conflit transcriptionnel entre deux unités de transcription très actives ne sera donc que marginal, voire complètement inexistant si la vitesse d'élongation est supérieure sur une matrice peu chromatinisée. Par contre, si l'assemblage du complexe de transcription est un processus lent ou s'il existe une pause dans l'élongation une transcription antisens peut effectivement entrer en conflit avec l'expression du gène, avec la nécessité qu'elle recouvre soit le promoteur soit le corps du gène en fonction de la nature du conflit. De nouveau, il est intéressant de remarquer que la taille des unités de transcription est probablement un paramètre essentiel dans le contexte de la régulation de l'expression. De fait, dans l'exemple le mieux caractérisé d'un conflit transcriptionnel, celui de *Igf2r* et *Airn*, les deux unités de transcription se recouvrent sur 30 kb. De plus, une version tronquée d'*Airn* qui ne recouvre que 16 kb d'*Igf2r* n'induit pas de silencing, ce qui est interprété comme indiquant que le silencing nécessite que la transcription antisens recouvre le promoteur.

Nos données obtenues avec la construction pStan sont à mettre en rapport avec les travaux de l'équipe de Nick Proudfoot (Gullerova et Proudfoot, 2012). Leur approche repose sur l'utilisation de constructions permettant une transcription convergente similaire à la nôtre (un gène entre deux promoteurs CMV) mais recherche un silencing en *trans* par un mécanisme transcriptionnel spécifique de séquence. Dans leur cas, la formation d'ARN double brin donne lieu à la formation de petits ARN qui peuvent agir sur le gène cellulaire aussi bien chez *S. Pombe* que chez les mammifères en cellules HeLa. *A contrario*, nos résultats montrent dans la même configuration un silencing en *cis* indépendant du mécanisme ARN interférant classique impliquant la protéine Ago2. Il est possible que la différence clé entre nos deux modèles soit la finalité recherchée, qui dans notre cas est un silencing en *cis* et dans le leur un silencing en *trans* sur un gène cellulaire. Ils s'abstiennent ainsi de toute recherche sur l'effet que peut avoir la construction sur elle-même et se contentent pour les possibles effets non spécifiques engendrés par l'expression transitoire d'une simple analyse de l'activation de la 2'5' oligoadenylate synthase qui est alors négative. D'autres études basées sur les travaux de Gullerova et Proudfoot montrent que chez le poisson zèbre, la transcription convergente du gène codant la protéine fluorescente mCherry permet un silencing efficace de cette dernière (Andrews *et al*, 2014). Ce silencing est alors Dicer dépendant et fait également appel à l'action de petit ARN comme c'est le cas



en *trans*. Cela soulève alors la question de savoir pourquoi notre modèle ne permet-il pas le même type de mécanisme ? Il est possible que les séquences de polyadénylation se trouvant en amont de chaque promoteur dans notre construction permettent un destin autre à ces transcrits. Par ailleurs dans nos travaux le devenir des transcrits n'a pas été étudié et il se peut que l'ARN double brin formé s'engage dans un mécanisme autre que l'ARN interférant. Des données récentes ont montré dans des cellules humaines que la formation d'un ARN double brin pouvait induire des activités biologiques indépendamment d'une activation de l'ARN interférence, lui permettant un rôle dans la régulation de la mitose (Portal *et al*, 2015). Un autre facteur est également à prendre en compte, dans les résultats de l'équipe de Proudfoot, la formation d'ARN double se fait dans un contexte en transitoire permettant peut être la formation de petits ARN. Du fait de la complexité de réponse en transitoire notre étude de la construction permettant la transcription convergente s'est faite essentiellement en clone stable. De cette intégration au génome peut résider la différence de mécanisme que nous observons. En effet, il est possible que l'intégration empêche la formation de siARN à partir de transcrit double brins dans notre cas. La grande majorité des cas de synthèse de siARN à partir de séquence sens/antisens endogène aussi appelé endo-siARN étant décrit dans la littérature dans les cellules de la lignée germinale (Werner *et al*, 2014 ; Tam *et al*, 2008).

Les longs ARN non codant interagissent fréquemment avec des complexes modificateurs de la chromatine ce qui peut permettre la mise en place d'un silencing capable de se maintenir après l'arrêt de la transcription antisens. C'est par exemple ce qui a été observé dans le cas de constructions dérivées de p15 et Anril (Yu *et al*, 2008). Lorsque j'ai introduit un site de recrutement pour PRC2 dans la construction pStan le silencing n'a pas perduré au-delà de l'expression de l'ARN antisens. Ceci peut refléter soit un déficit dans le recrutement de PRC2, puisqu'ici il n'a pas été mesuré directement par des expériences d'immunoprécipitation de la chromatine, soit que l'organisation de la construction n'est pas permissive pour une telle régulation. On peut d'ailleurs remarquer que le site de recrutement de PRC2 est placé à mi-distance entre les deux promoteurs de la construction, lesquels ne sont distants que de 1, 5 kb. Dans ces conditions, on pourrait plutôt s'attendre à la mise en place d'un silencing qui affecte aussi bien la transcription sens que la transcription antisens. Une population de cellules « double négatives » peut d'ailleurs être observée dans les clones pStan après une





exposition prolongée à la doxycycline. Cependant cette population apparaît aussi dans les cellules maintenues en parallèle en l'absence de doxycycline et semble donc plutôt être liée à la mise en place d'un silencing au niveau du site d'intégration. L'expression de l'ARN antisens n'accélère pas vraiment la mise en place de ce silencing.

### 3 - Bilan et Perspectives

Les études en transitoire soulèvent deux questions principales. D'une part, quel est le mécanisme du silencing non spécifique induit par l'expression d'un ARN antisens dans nos expériences ? D'autre part, comment sont distingués les transcrits d'origine épisomique ou génomique ?

Pour le mécanisme du silencing, j'ai testé le rôle de PKR et de la RNaseL sans observer d'effet marqué. Des expériences préliminaires ne montrent pas non plus d'effet important suite au silencing de RIG-I et MDA5 (données non présentées). L'explication la plus simple est que nous induisons une activation de l'ensemble de la voie interféron et donc que le silencing de chacun de ces acteurs séparément ne donne lieu qu'à une faible réponse. Une étude détaillée du niveau d'activation de chacune de ces voies permettrait de valider cette hypothèse tandis que leur silencing simultané devrait conduire à un effet plus marqué. Il reste néanmoins à comprendre pourquoi dans l'étude du groupe de P. Svoboda, la kinase PKR semble à elle-seule jouer un rôle déterminant dans le silencing non spécifique. Ce d'autant plus que les approches expérimentales sont très voisines, transfection au phosphate de calcium en cellules HEK 293, entre autres. On peut concevoir que la structure en tige-boucle présente en aval d'une région codante puisse être un meilleur activateur de PKR qu'un ARN double brin non traductible. Cependant, la longueur de la région tige-boucle introduite, 520 pb n'est pas sensiblement différente de celle des duplex qui peuvent se former dans nos expériences (~ 750 pb). Il faudrait donc supposer qu'en fonction du contexte, les ARN double brin sont de meilleurs activateurs des différentes classes de détecteurs d'ARN double brin présents dans la cellule et que nos duplex sont plus efficaces pour activer les autres voies que PKR.

L'observation la plus frappante de cette partie du travail est cependant la différence de sensibilité des transcrits au silencing non spécifique en fonction de leur origine épisomique



ou génomique. Cette observation n'est pas nouvelle (Terenzi *et al*, 1999) et a été renouvelée par le groupe de P. Svoboda mais n'a pas fait l'objet d'études plus approfondies pour en définir les bases moléculaires. D'un côté, une telle distinction s'intègre de manière cohérente dans le contexte biologique d'une réponse antivirale. Si, au moins dans un premier temps, la cellule peut cibler spécifiquement les ARN « extérieurs au génome » ceci devrait permettre d'atténuer l'infection sans compromettre la viabilité cellulaire. D'un autre côté, une telle distinction suppose qu'il existe une marque spécifique de l'une ou l'autre classe de transcrits qui s'ajouterait ainsi aux signatures moléculaires des pathogènes. L'identification de cette marque constituerait un prolongement évident de cette partie de mon travail. Les données présentées montrent que la différence de sensibilité est très marquée entre les deux classes de transcrits et au moins en première analyse on peut s'attendre à ce que la marque moléculaire soit portée par une majorité de transcrits. Un premier élément d'orientation serait de rechercher dans quel compartiment cellulaire a lieu le silencing ; puisque les mécanismes évoqués appartiennent à la voie de l'interféron et sont principalement cytoplasmiques il a été implicitement supposé que le silencing non spécifique était lui aussi cytoplasmique mais ceci reste à valider. Des trois pistes évoquées plus haut, celle de la queue polyA est la plus facile à tester mais aussi la moins vraisemblable puisque les transcrits génomiques ont une distribution très large de la longueur de leur queue polyA. L'hypothèse d'une localisation cellulaire spécifique est assez facile à envisager dans le noyau mais devrait *a priori* être maintenue dans le cytoplasme si le silencing y a lieu. Aussi une différence dans la couverture protéique du transcrit apparaît comme l'hypothèse la plus plausible. Une purification des transcrits exprimés à partir d'un plasmide en transitoire ou après intégration dans le génome suivie d'une analyse en spectrométrie de masse paraît donc s'imposer ; celle-ci pourrait être facilitée par l'incorporation dans le transcrit d'un motif permettant sa purification, des sites MS2 par exemple.

Pour les expériences en clones stables, le premier message est la faiblesse de la régulation par les transcrits antisens. Si ceci n'était pas inattendu pour un transcrit exprimé en *trans* au vu du faible nombre d'exemple décrits, il est souvent supposé que l'expression en *cis* devrait être un mécanisme majeur de silencing. Or, même si nous avons observé un silencing dans les clones stables pStan, l'ampleur du silencing reste modeste. De plus, il n'existe pas de lien clair entre le niveau d'expression du transcrit antisens et l'amplitude du silencing puisque, de



manière globale, les forts expresseurs antisens sont aussi les forts expresseurs sens. Une limitation déjà évoquée de l'approche utilisée est la proximité entre les deux promoteurs sens et antisens dans nos constructions. Néanmoins, la co-expression que nous observons indique que le conflit transcriptionnel n'est pas en soi un obstacle majeur à l'expression et, ce, même lorsque les deux unités de transcription sont très fortement exprimées. Cette conclusion est renforcée par l'absence de corrélation négative entre l'expression des deux gènes dans les expériences en transitoire. De même, on peut conclure que la possibilité de former un duplex d'ARN n'induit pas non plus un silencing très fort, toujours dans des conditions de forts niveaux d'expression. Ceci est à replacer dans le contexte des données génomiques où le transcrit antisens est souvent exprimé à des niveaux nettement plus faibles que le transcrit sens. Ainsi, il apparaît qu'en l'absence de mécanisme spécifique comme le recrutement de complexes modificateurs de la chromatine, la transcription antisens n'est pas un obstacle majeur à l'expression d'un gène même si elle est susceptible d'en moduler l'expression. De fait, l'hypothèse retenue pour expliquer le silencing dans les clones stables pStan est celle d'un conflit transcriptionnel par la transcription antisens qui recouvre le promoteur mais il resterait à établir ce mécanisme. La difficulté vient de notre connaissance partielle de la terminaison de la transcription. La stratégie utilisée habituellement d'introduire un site de polyadénylation n'est évidemment qu'un moyen indirect de provoquer la terminaison de la transcription en fonction des caractéristiques de l'élongation au-delà de ce site. Aussi pour exploiter cette stratégie il conviendrait d'éloigner les deux promoteurs de la construction afin de laisser suffisamment de séquence en aval du site de polyadénylation pour que la transcription s'arrête avant de recouvrir le promoteur. Il est aussi possible que le choix du promoteur CMV soit en partie à l'origine de la faiblesse des effets observés. En effet, ce promoteur fort d'origine virale est peut-être plus résistant à de telles régulations que la majorité des promoteurs cellulaires. De ce point de vue l'utilisation de deux promoteurs régulés indépendamment et dont le niveau d'expression soit contrôlable de manière graduelle pourrait permettre d'explorer dans quelles conditions d'expression un silencing peut se mettre en place.



## Bibliographie

- Abdelmohsen, K. & Gorospe, M. (2010) Posttranscriptional regulation of cancer traits by HuR. *Wiley Interdiscip Rev RNA*, **1**, 214-229.
- Alexopoulou, L., Holt, A. C., Medzhitov, R. & Flavell, R. A. (2001) Recognition of double-stranded RNA and activation of NF-kappaB by Toll-like receptor 3. *Nature*, **413**, 732-738.
- Altuvia, Y., Landgraf, P., Lithwick, G., Elefant, N., Pfeffer, S., Aravin, A., *et al.* (2005) Clustering and conservation patterns of human microRNAs. *Nucleic Acids Res*, **33**, 2697-2706.
- Andrews, O. E., Cha, D. J., Wei, C. & Patton, J. G. (2014) RNAi-mediated gene silencing in zebrafish triggered by convergent transcription. *Sci Rep*, **4**, 5222.
- Aravin, A., Gaidatzis, D., Pfeffer, S., Lagos-Quintana, M., Landgraf, P., Iovino, N., *et al.* (2006) A novel class of small RNAs bind to MILI protein in mouse testes. *Nature*, **442**, 203-207.
- Aravin, A. A., Lagos-Quintana, M., Yalcin, A., Zavolan, M., Marks, D., Snyder, B., *et al.* (2003) The small RNA profile during *Drosophila melanogaster* development. *Dev Cell*, **5**, 337-350.
- Aravin, A. A., Naumova, N. M., Tulin, A. V., Vagin, V. V., Rozovsky, Y. M. & Gvozdev, V. A. (2001) Double-stranded RNA-mediated silencing of genomic tandem repeats and transposable elements in the *D. melanogaster* germline. *Curr Biol*, **11**, 1017-1027.
- Aravin, A. A., Sachidanandam, R., Bourc'his, D., Schaefer, C., Pezic, D., Toth, K. F., *et al.* (2008) A piRNA pathway primed by individual transposons is linked to de novo DNA methylation in mice. *Mol Cell*, **31**, 785-799.
- Arkov, A. L., Wang, J. Y., Ramos, A. & Lehmann, R. (2006) The role of Tudor domains in germline development and polar granule architecture. *Development*, **133**, 4053-4062.
- Bagga, S., Bracht, J., Hunter, S., Massirer, K., Holtz, J., Eachus, R., *et al.* (2005) Regulation by let-7 and lin-4 miRNAs results in target mRNA degradation. *Cell*, **122**, 553-563.
- Banerjee, D. & Slack, F. (2002) Control of developmental timing by small temporal RNAs: a paradigm for RNA-mediated regulation of gene expression. *Bioessays*, **24**, 119-129.
- Basilio C, Wahba AJ, Lengyel P, Speyer JF, Ochoa S. (1962) Synthetic polynucleotides and the amino acid code. *Proc Natl Acad Sci USA*, **48**, 613-6.
- Batzer, M. A. & Deininger, P. L. (2002) Alu repeats and human genomic diversity. *Nat Rev Genet*, **3**, 370-379.
- Bazak, L. Haviv, A. Barak, M. Jacob-Hirsh, J. Deng, P *et al.* (2013) A-to-I RNA editing occurs at over a hundred million genomic sites, located in majority of human genes. *Genome Res*, **24**, 365-376.
- Behm-Ansmant, I., Rehwinkel, J., Doerks, T., Stark, A., Bork, P. & Izaurralde, E. (2006) mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev*, **20**, 1885-1898.



- Beltran, M., Puig, I., Pena, C., Garcia, J. M., Alvarez, A. B., Pena, R., *et al.* (2008) A natural antisense transcript regulates Zeb2/Sip1 gene expression during Snail1-induced epithelial-mesenchymal transition. *Genes Dev*, **22**, 756-769.
- Ben-Asouli, Y., Banai, Y., Pel-Or, Y., Shir, A. & Kaempfer, R. (2002) Human interferon-gamma mRNA autoregulates its translation through a pseudoknot that activates the interferon-inducible protein kinase PKR. *Cell*, **108**, 221-232.
- Berghoff, E. G., Clark, M. F., Chen, S., Cajigas, I., Leib, D. E. & Kohtz, J. D. (2013) Evf2 (Dlx6as) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. *Development*, **140**, 4407-4416.
- Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. (2001) Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature*, **409**, 363-366.
- Bevilacqua, P. C. & Cech, T. R. (1996) Minor-groove recognition of double-stranded RNA by the double-stranded RNA-binding domain from the RNA-activated protein kinase PKR. *Biochemistry*, **35**, 9983-9994.
- Blasius, A. L. & Beutler, B. (2010) Intracellular toll-like receptors. *Immunity*, **32**, 305-315.
- Blencowe, B. J., Bauren, G., Eldridge, A. G., Issner, R., Nickerson, J. A., Rosonina, E., *et al.* (2000) The SRm160/300 splicing coactivator subunits. *RNA*, **6**, 111-120.
- Bohmert, K., Camus, I., Bellini, C., Bouchez, D., Caboche, M. & Benning, C. (1998) AGO1 defines a novel locus of Arabidopsis controlling leaf development. *EMBO J*, **17**, 170-180.
- Bond, A. M., Vangompel, M. J., Sametsky, E. A., Clark, M. F., Savage, J. C., Disterhoft, J. F., *et al.* (2009) Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. *Nat Neurosci*, **12**, 1020-1027.
- Borchert, G. M., Lanier, W. & Davidson, B. L. (2006) RNA polymerase III transcribes human microRNAs. *Nat Struct Mol Biol*, **13**, 1097-1101.
- Borsani, G., Tonlorenzi, R., Simmler, M. C., Dandolo, L., Arnaud, D., Capra, V., *et al.* (1991) Characterization of a murine gene expressed from the inactive X chromosome. *Nature*, **351**, 325-329.
- Bourgeois, C. F., Lejeune, F. & Stevenin, J. (2004) Broad specificity of SR (serine/arginine) proteins in the regulation of alternative splicing of pre-messenger RNA. *Prog Nucleic Acid Res Mol Biol*, **78**, 37-88.
- Brennecke, J., Aravin, A. A., Stark, A., Dus, M., Kellis, M., Sachidanandam, R., *et al.* (2007) Discrete small RNA-generating loci as master regulators of transposon activity in *Drosophila*. *Cell*, **128**, 1089-1103.
- Brockdorff, N., Ashworth, A., Kay, G. F., McCabe, V. M., Norris, D. P., Cooper, P. J., *et al.* (1992) The product of the mouse Xist gene is a 15 kb inactive X-specific transcript containing no conserved ORF and located in the nucleus. *Cell*, **71**, 515-526.
- Brown, C. J., Lafreniere, R. G., Powers, V. E., Sebastio, G., Ballabio, A., Pettigrew, A. L., *et al.* (1991) Localization of the X inactivation centre on the human X chromosome in Xq13. *Nature*, **349**, 82-84.
- Buhler, M., Verdel, A. & Moazed, D. (2006) Tethering RITS to a nascent transcript initiates RNAi- and heterochromatin-dependent gene silencing. *Cell*, **125**, 873-886.



- Burd, C. E., Jeck, W. R., Liu, Y., Sanoff, H. K., Wang, Z. & Sharpless, N. E. (2010) Expression of linear and novel circular forms of an INK4/ARF-associated non-coding RNA correlates with atherosclerosis risk. *PLoS Genet*, **6**, e1001233.
- Cao, W., Jamison, S. F. & Garcia-Blanco, M. A. (1997) Both phosphorylation and dephosphorylation of ASF/SF2 are required for pre-mRNA splicing in vitro. *RNA*, **3**, 1456-1467.
- Carninci, P., Kasukawa, T., Katayama, S., Gough, J., Frith, M. C., Maeda, N., *et al.* (2005) The transcriptional landscape of the mammalian genome. *Science*, **309**, 1559-1563.
- Casa, V. & Gabellini, D. (2012) A repetitive elements perspective in Polycomb epigenetics. *Front Genet*, **3**, 199.
- Chakrabarti, A., Jha, B. K. & Silverman, R. H. (2011) New insights into the role of RNase L in innate immunity. *J Interferon Cytokine Res*, **31**, 49-57.
- Chen, C. Y., Zheng, D., Xia, Z. & Shyu, A. B. (2009) Ago-TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat Struct Mol Biol*, **16**, 1160-1166.
- Chrousos, G. P. & Kino, T. (2005) Intracellular glucocorticoid signaling: a formerly simple system turns stochastic. *Sci STKE*, **2005**, pe48.
- Chu, C. Y. & Rana, T. M. (2006) Translation repression in human cells by microRNA-induced gene silencing requires RCK/p54. *PLoS Biol*, **4**, e210.
- Cohen-Chalamish, S., Hasson, A., Weinberg, D., Namer, L. S., Banai, Y., Osman, F., *et al.* (2009) Dynamic refolding of IFN-gamma mRNA enables it to function as PKR activator and translation template. *Nat Chem Biol*, **5**, 896-903.
- Coller, J. & Parker, R. (2005) General translational repression by activators of mRNA decapping. *Cell*, **122**, 875-886.
- Conaway, J. W., Shilatifard, A., Dvir, A. & Conaway, R. C. (2000) Control of elongation by RNA polymerase II. *Trends Biochem Sci*, **25**, 375-380.
- Conley, A. B. & Jordan, I. K. (2012) Cell type-specific termination of transcription by transposable element sequences. *Mob DNA*, **3**, 15.
- Conley, A. B., Miller, W. J. & Jordan, I. K. (2008) Human cis natural antisense transcripts initiated by transposable elements. *Trends Genet*, **24**, 53-56.
- Denholtz, M., Bonora, G., Chronis, C., Splinter, E., de Laat, W., Ernst, J., *et al.* (2013) Long-range chromatin contacts in embryonic stem cells reveal a role for pluripotency factors and polycomb proteins in genome organization. *Cell Stem Cell*, **13**, 602-616.
- Denli, A. M., Tops, B. B., Plasterk, R. H., Ketting, R. F. & Hannon, G. J. (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature*, **432**, 231-235.
- Di Cristo, G. (2007) Development of cortical GABAergic circuits and its implications for neurodevelopmental disorders. *Clin Genet*, **72**, 1-8.
- Ding, S. W. & Voinnet, O. (2007) Antiviral immunity directed by small RNAs. *Cell*, **130**, 413-426.





- Dinger, M. E., Pang, K. C., Mercer, T. R. & Mattick, J. S. (2008) Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput Biol*, **4**, e1000176.
- Drinnenberg, I. A., Fink, G. R. & Bartel, D. P. (2011) Compatibility with killer explains the rise of RNAi-deficient fungi. *Science*, **333**, 1592.
- Duret, L., Chureau, C., Samain, S., Weissenbach, J. & Avner, P. (2006) The Xist RNA gene evolved in eutherians by pseudogenization of a protein-coding gene. *Science*, **312**, 1653-1655.
- Eichhorn, S. W., Guo, H., McGeary, S. E., Rodriguez-Mias, R. A., Shin, C., Baek, D., *et al.* (2014) mRNA destabilization is the dominant effect of mammalian microRNAs by the time substantial repression ensues. *Mol Cell*, **56**, 104-115.
- Elbashir, S. M., Harborth, J., Lendeckel, W., Yalcin, A., Weber, K. & Tuschl, T. (2001) Duplexes of 21-nucleotide RNAs mediate RNA interference in cultured mammalian cells. *Nature*, **411**, 494-498.
- Elisaphenko, E. A., Kolesnikov, N. N., Shevchenko, A. I., Rogozin, I. B., Nesterova, T. B., Brockdorff, N., *et al.* (2008) A dual origin of the Xist gene from a protein-coding gene and a set of transposable elements. *PLoS One*, **3**, e2521.
- Eystathiou, T., Jakymiw, A., Chan, E. K., Seraphin, B., Cougot, N. & Fritzler, M. J. (2003) The GW182 protein colocalizes with mRNA degradation associated proteins hDcp1 and hLSm4 in cytoplasmic GW bodies. *RNA*, **9**, 1171-1173.
- Faghihi, M. A., Modarresi, F., Khalil, A. M., Wood, D. E., Sahagan, B. G., Morgan, T. E., *et al.* (2008) Expression of a noncoding RNA is elevated in Alzheimer's disease and drives rapid feed-forward regulation of beta-secretase. *Nat Med*, **14**, 723-730.
- Faghihi, M. A. & Wahlestedt, C. (2006) RNA interference is not involved in natural antisense mediated regulation of gene expression in mammals. *Genome Biol*, **7**, R38.
- Faghihi, M. A., Zhang, M., Huang, J., Modarresi, F., Van der Brug, M. P., Nalls, M. A., *et al.* (2010) Evidence for natural antisense transcript-mediated inhibition of microRNA function. *Genome Biol*, **11**, R56.
- Feng, J., Bi, C., Clark, B. S., Mady, R., Shah, P. & Kohtz, J. D. (2006) The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev*, **20**, 1470-1484.
- Findley, S. D., Tamanaha, M., Clegg, N. J. & Ruohola-Baker, H. (2003) Maelstrom, a *Drosophila* spindle-class gene, encodes a protein that colocalizes with Vasa and RDE1/AGO1 homolog, Aubergine, in nuage. *Development*, **130**, 859-871.
- Fire, A., Xu, S., Montgomery, M. K., Kostas, S. A., Driver, S. E. & Mello, C. C. (1998) Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature*, **391**, 806-811.
- Floyd-Smith, G., Slattery, E. & Lengyel, P. (1981) Interferon action: RNA cleavage pattern of a (2'-5')oligoadenylate--dependent endonuclease. *Science*, **212**, 1030-1032.
- Folkersen, L., Kyriakou, T., Goel, A., Peden, J., Malarstig, A., Paulsson-Berne, G., *et al.* (2009) Relationship between CAD risk genotype in the chromosome 9p21 locus and gene expression. Identification of eight new ANRIL splice variants. *PLoS One*, **4**, e7677.



- Gack, M. U., Shin, Y. C., Joo, C. H., Urano, T., Liang, C., Sun, L., *et al.* (2007) TRIM25 RING-finger E3 ubiquitin ligase is essential for RIG-I-mediated antiviral activity. *Nature*, **446**, 916-920.
- Garcia, M. A., Gil, J., Ventoso, I., Guerra, S., Domingo, E., Rivas, C., *et al.* (2006) Impact of protein kinase PKR in cell biology: from antiviral to antiproliferative action. *Microbiol Mol Biol Rev*, **70**, 1032-1060.
- Garcia, S., Billecocq, A., Crance, J. M., Prins, M., Garin, D. & Bouloy, M. (2006) Viral suppressors of RNA interference impair RNA silencing induced by a Semliki Forest virus replicon in tick cells. *J Gen Virol*, **87**, 1985-1989.
- Gardner, E. J., Nizami, Z. F., Talbot, C. C., Jr. & Gall, J. G. (2012) Stable intronic sequence RNA (sisRNA), a new class of noncoding RNA from the oocyte nucleus of *Xenopus tropicalis*. *Genes Dev*, **26**, 2550-2559.
- Garrett, S. Rosenthal, JJ. (2012) RNA editing underlies temperature adaptation in K<sup>+</sup> channels from polar octopuses. *Science*, **335**, 848-851.
- Geng, Y. J., Xie, S. L., Li, Q., Ma, J. & Wang, G. Y. (2011) Large intervening non-coding RNA HOTAIR is associated with hepatocellular carcinoma progression. *J Int Med Res*, **39**, 2119-2128.
- Girard, A., Sachidanandam, R., Hannon, G. J. & Carmell, M. A. (2006) A germline-specific class of small RNAs binds mammalian Piwi proteins. *Nature*, **442**, 199-202.
- Goh, W. S., Falciatori, I., Tam, O. H., Burgess, R., Meikar, O., Kotaja, N., *et al.* (2015) piRNA-directed cleavage of meiotic transcripts regulates spermatogenesis. *Genes Dev*, **29**, 1032-1044.
- Gong, C. & Maquat, L. E. (2011) "Alu"strious long ncRNAs and their role in shortening mRNA half-lives. *Cell Cycle*, **10**, 1882-1883.
- Gong, C., Popp, M. W. & Maquat, L. E. (2012) Biochemical analysis of long non-coding RNA-containing ribonucleoprotein complexes. *Methods*, **58**, 88-93.
- Gregory, R. I., Yan, K. P., Amuthan, G., Chendrimada, T., Doratotaj, B., Cooch, N., *et al.* (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature*, **432**, 235-240.
- Gullerova, M. & Proudfoot, N. J. (2012) Convergent transcription induces transcriptional gene silencing in fission yeast and mammalian cells. *Nat Struct Mol Biol*, **19**, 1193-1201.
- Gunawardane, L. S., Saito, K., Nishida, K. M., Miyoshi, K., Kawamura, Y., Nagami, T., *et al.* (2007) A slicer-mediated mechanism for repeat-associated siRNA 5' end formation in *Drosophila*. *Science*, **315**, 1587-1590.
- Guo, L. & Lu, Z. (2010) The fate of miRNA\* strand through evolutionary analysis: implication for degradation as merely carrier strand or potential regulatory molecule? *PLoS One*, **5**, e11387.
- Gupta, R. A., Shah, N., Wang, K. C., Kim, J., Horlings, H. M., Wong, D. J., *et al.* (2010) Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature*, **464**, 1071-1076.





Guttman, M., Amit, I., Garber, M., French, C., Lin, M. F., Feldser, D., *et al.* (2009) Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature*, **458**, 223-227.

Guttman, M., Donaghey, J., Carey, B. W., Garber, M., Grenier, J. K., Munson, G., *et al.* (2011) lincRNAs act in the circuitry controlling pluripotency and differentiation. *Nature*, **477**, 295-300.

Guttman, M., Russell, P., Ingolia, N. T., Weissman, J. S. & Lander, E. S. (2013) Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell*, **154**, 240-251.

Haase, A. D., Fenoglio, S., Muerdter, F., Guzzardo, P. M., Czech, B., Pappin, D. J., *et al.* (2010) Probing the initiation and effector phases of the somatic piRNA pathway in *Drosophila*. *Genes Dev*, **24**, 2499-2504.

Hamilton, A. J. & Baulcombe, D. C. (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science*, **286**, 950-952.

Hammond, S. M., Bernstein, E., Beach, D. & Hannon, G. J. (2000) An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells. *Nature*, **404**, 293-296.

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., *et al.* (2004) Species-specific recognition of single-stranded RNA via toll-like receptor 7 and 8. *Science*, **303**, 1526-1529.

Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., *et al.* (2000) A Toll-like receptor recognizes bacterial DNA. *Nature*, **408**, 740-745.

Higuchi, M. Single, FN. Kohler, M. Sommer, B. Sprengel, R. Seeburg, PH. (1993) RNA editing of AMPA receptor subunit GluR-B: A base-paired intron-exon structure determines position and efficiency. *Cell*, **75**, 1361-1370.

Hinman, M. N. & Lou, H. (2008) Diverse molecular functions of Hu proteins. *Cell Mol Life Sci*, **65**, 3168-3181.

Hoagland, M. B., Stephenson, M. L., Scott, J. F., Hecht, L. I. & Zamecnik, P. C. (1958) A soluble ribonucleic acid intermediate in protein synthesis. *J Biol Chem*, **231**, 241-257.

Horwich, M. D., Li, C., Matranga, C., Vagin, V., Farley, G., Wang, P., *et al.* (2007) The *Drosophila* RNA methyltransferase, DmHen1, modifies germline piRNAs and single-stranded siRNAs in RISC. *Curr Biol*, **17**, 1265-1272.

Hovanessian, A. G., Laurent, A. G., Chebath, J., Galabru, J., Robert, N. & Svab, J. (1987) Identification of 69-kd and 100-kd forms of 2-5A synthetase in interferon-treated human cells by specific monoclonal antibodies. *EMBO J*, **6**, 1273-1280.

Huarte, M., Guttman, M., Feldser, D., Garber, M., Koziol, M. J., Kenzelmann-Broz, D., *et al.* (2010) A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell*, **142**, 409-419.

Ipsaro, J. J., Haase, A. D., Knott, S. R., Joshua-Tor, L. & Hannon, G. J. (2012) The structural biochemistry of Zucchini implicates it as a nuclease in piRNA biogenesis. *Nature*, **491**, 279-283.



Jammi, N. V., Whitby, L. R. & Beal, P. A. (2003) Small molecule inhibitors of the RNA-dependent protein kinase. *Biochem Biophys Res Commun*, **308**, 50-57.

Jeon, Y. & Lee, J. T. (2011) YY1 tethers Xist RNA to the inactive X nucleation center. *Cell*, **146**, 119-133.

Ji, P., Diederichs, S., Wang, W., Boing, S., Metzger, R., Schneider, P. M., *et al.* (2003) MALAT-1, a novel noncoding RNA, and thymosin beta4 predict metastasis and survival in early-stage non-small cell lung cancer. *Oncogene*, **22**, 8031-8041.

Johnson, J. M., Edwards, S., Shoemaker, D. & Schadt, E. E. (2005) Dark matter in the genome: evidence of widespread transcription detected by microarray tiling experiments. *Trends Genet*, **21**, 93-102.

Johnsson, P., Ackley, A., Vidarsdottir, L., Lui, W. O., Corcoran, M., Grander, D., *et al.* (2013) A pseudogene long-noncoding-RNA network regulates PTEN transcription and translation in human cells. *Nat Struct Mol Biol*, **20**, 440-446.

Källman, AM. Sahlin, M. Ohman, M. (2003) ADAR2 A→I editing: site selectivity and editing efficiency are separate events., *Nucleic Acids Res*, **31**, 4874-4881.

Kaneko, S., Li, G., Son, J., Xu, C. F., Margueron, R., Neubert, T. A., *et al.* (2010) Phosphorylation of the PRC2 component Ezh2 is cell cycle-regulated and up-regulates its binding to ncRNA. *Genes Dev*, **24**, 2615-2620.

Kapranov, P., Drenkow, J., Cheng, J., Long, J., Helt, G., Dike, S., *et al.* (2005) Examples of the complex architecture of the human transcriptome revealed by RACE and high-density tiling arrays. *Genome Res*, **15**, 987-997.

Katayama, S., Tomaru, Y., Kasukawa, T., Waki, K., Nakanishi, M., Nakamura, M., *et al.* (2005) Antisense transcription in the mammalian transcriptome. *Science*, **309**, 1564-1566.

Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., *et al.* (2008) Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. *J Exp Med*, **205**, 1601-1610.

Kawaoka, S., Izumi, N., Katsuma, S. & Tomari, Y. (2011) 3' end formation of PIWI-interacting RNAs in vitro. *Mol Cell*, **43**, 1015-1022.

Kennerdell, J. R. & Carthew, R. W. (1998) Use of dsRNA-mediated genetic interference to demonstrate that frizzled and frizzled 2 act in the wingless pathway. *Cell*, **95**, 1017-1026.

Khalil, A. M., Guttman, M., Huarte, M., Garber, M., Raj, A., Rivea Morales, D., *et al.* (2009) Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A*, **106**, 11667-11672.

Kim, H. H., Kuwano, Y., Srikantan, S., Lee, E. K., Martindale, J. L. & Gorospe, M. (2009) HuR recruits let-7/RISC to repress c-Myc expression. *Genes Dev*, **23**, 1743-1748.

Kim, U. Garner, TL. Sanford, T. Speicher, D. Murray, JM. Nishikura, K. (1994) Purification and characterization of double-stranded RNA adenosine deaminase from bovine nuclear extracts. *J Biol Chem*, **269**, 13480-13489.

Kim, Y. K., Furic, L., Desgroseillers, L. & Maquat, L. E. (2005) Mammalian Staufen1 recruits Upf1 to specific mRNA 3'UTRs so as to elicit mRNA decay. *Cell*, **120**, 195-208.



- Kim, Y. K., Furic, L., Parisien, M., Major, F., DesGroseillers, L. & Maquat, L. E. (2007) Staufen1 regulates diverse classes of mammalian transcripts. *EMBO J*, **26**, 2670-2681.
- Kino, T., Hurt, D. E., Ichijo, T., Nader, N. & Chrousos, G. P. (2010) Noncoding RNA gas5 is a growth arrest- and starvation-associated repressor of the glucocorticoid receptor. *Sci Signal*, **3**, ra8.
- Kotake, Y., Nakagawa, T., Kitagawa, K., Suzuki, S., Liu, N., Kitagawa, M., *et al.* (2011) Long non-coding RNA ANRIL is required for the PRC2 recruitment to and silencing of p15(INK4B) tumor suppressor gene. *Oncogene*, **30**, 1956-1962.
- Kotelnikov, R. N., Klenov, M. S., Rozovsky, Y. M., Olenina, L. V., Kibanov, M. V. & Gvozdev, V. A. (2009) Peculiarities of piRNA-mediated post-transcriptional silencing of Stellate repeats in testes of *Drosophila melanogaster*. *Nucleic Acids Res*, **37**, 3254-3263.
- Kowalik, K. M., Shimada, Y., Flury, V., Stadler, M. B., Batki, J. & Buhler, M. (2015) The Paf1 complex represses small-RNA-mediated epigenetic gene silencing. *Nature*, **520**, 248-252.
- Kristiansen, H., Gad, H. H., Eskildsen-Larsen, S., Despres, P. & Hartmann, R. (2011) The oligoadenylate synthetase family: an ancient protein family with multiple antiviral activities. *J Interferon Cytokine Res*, **31**, 41-47.
- Kuramochi-Miyagawa, S., Kimura, T., Yomogida, K., Kuroiwa, A., Tadokoro, Y., Fujita, Y., *et al.* (2001) Two mouse piwi-related genes: miwi and mili. *Mech Dev*, **108**, 121-133.
- Lagos-Quintana, M., Rauhut, R., Lendeckel, W. & Tuschl, T. (2001) Identification of novel genes coding for small expressed RNAs. *Science*, **294**, 853-858.
- Lai, E. C., Tomancak, P., Williams, R. W. & Rubin, G. M. (2003) Computational identification of *Drosophila* microRNA genes. *Genome Biol*, **4**, R42.
- Latos, P. A., Pauler, F. M., Koerner, M. V., Senergin, H. B., Hudson, Q. J., Stocsits, R. R., *et al.* (2012) Airn transcriptional overlap, but not its lncRNA products, induces imprinted Igf2r silencing. *Science*, **338**, 1469-1472.
- Lazzaretti, D., Tournier, I. & Izaurralde, E. (2009) The C-terminal domains of human TNRC6A, TNRC6B, and TNRC6C silence bound transcripts independently of Argonaute proteins. *RNA*, **15**, 1059-1066.
- Le Thomas, A., Rogers, A. K., Webster, A., Marinov, G. K., Liao, S. E., Perkins, E. M., *et al.* (2013) Piwi induces piRNA-guided transcriptional silencing and establishment of a repressive chromatin state. *Genes Dev*, **27**, 390-399.
- Lee, R. C., Feinbaum, R. L. & Ambros, V. (1993) The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, **75**, 843-854.
- Lee, Y., Jeon, K., Lee, J. T., Kim, S. & Kim, V. N. (2002) MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J*, **21**, 4663-4670.
- Leung, A. K. & Sharp, P. A. (2013) Quantifying Argonaute proteins in and out of GW/P-bodies: implications in microRNA activities. *Adv Exp Med Biol*, **768**, 165-182.
- Li, S., Lian, S. L., Moser, J. J., Fritzler, M. L., Fritzler, M. J., Satoh, M., *et al.* (2008) Identification of GW182 and its novel isoform TNGW1 as translational repressors in Ago2-mediated silencing. *J Cell Sci*, **121**, 4134-4144.



- Lim, L. P., Glasner, M. E., Yekta, S., Burge, C. B. & Bartel, D. P. (2003) Vertebrate microRNA genes. *Science*, **299**, 1540.
- Lim, L. P., Lau, N. C., Weinstein, E. G., Abdelhakim, A., Yekta, S., Rhoades, M. W., *et al.* (2003) The microRNAs of *Caenorhabditis elegans*. *Genes Dev*, **17**, 991-1008.
- Lin, R., Maeda, S., Liu, C., Karin, M. & Edgington, T. S. (2007) A large noncoding RNA is a marker for murine hepatocellular carcinomas and a spectrum of human carcinomas. *Oncogene*, **26**, 851-858.
- Lin, S. & Fu, X. D. (2007) SR proteins and related factors in alternative splicing. *Adv Exp Med Biol*, **623**, 107-122.
- Liu, J., Valencia-Sanchez, M. A., Hannon, G. J. & Parker, R. (2005) MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol*, **7**, 719-723.
- Long, J. C. & Cáceres, J. F. (2009) The SR protein family of splicing factors: master regulators of gene expression. *Biochem J*, **417**, 15-27.
- Lund, E., Guttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. (2004) Nuclear export of microRNA precursors. *Science*, **303**, 95-98.
- Lykke-Andersen, J. (2002) Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol*, **22**, 8114-8121.
- Lyle, R., Watanabe, D., te Vrugte, D., Lerchner, W., Smrzka, O. W., Wutz, A., *et al.* (2000) The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat Genet*, **25**, 19-21.
- Mahmoudi, S., Henriksson, S., Corcoran, M., Mendez-Vidal, C., Wiman, K. G. & Farnebo, M. (2009) Wrap53, a natural p53 antisense transcript required for p53 induction upon DNA damage. *Mol Cell*, **33**, 462-471.
- Manche, L., Green, S. R., Schmedt, C. & Mathews, M. B. (1992) Interactions between double-stranded RNA regulators and the protein kinase DAI. *Mol Cell Biol*, **12**, 5238-5248.
- Marcucci, R., Brindle, J., Paro, S., Casadio, A., Hempel, S., Morrice, N. *et al.* (2011) Pin1 and WWP2 regulate GluR2 Q/R site RNA editing by ADAR2 with opposing effects. *Embo J*, **30**, 4211-4222.
- Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. (2005) Passenger-strand cleavage facilitates assembly of siRNA into Ago2-containing RNAi enzyme complexes. *Cell*, **123**, 607-620.
- McKenna, S. A., Lindhout, D. A., Kim, I., Liu, C. W., Gelev, V. M., Wagner, G., *et al.* (2007) Molecular framework for the activation of RNA-dependent protein kinase. *J Biol Chem*, **282**, 11474-11486.
- Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F. & Dietz, H. C. (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet*, **36**, 1073-1078.



- Mercer, T. R., Dinger, M. E., Bracken, C. P., Kolle, G., Szubert, J. M., Korbie, D. J., *et al.* (2010) Regulated post-transcriptional RNA cleavage diversifies the eukaryotic transcriptome. *Genome Res*, **20**, 1639-1650.
- Mercer, T. R., Dinger, M. E., Sunkin, S. M., Mehler, M. F. & Mattick, J. S. (2008) Specific expression of long noncoding RNAs in the mouse brain. *Proc Natl Acad Sci U S A*, **105**, 716-721.
- Mohn, F., Handler, D. & Brennecke, J. (2015) Noncoding RNA. piRNA-guided slicing specifies transcripts for Zucchini-dependent, phased piRNA biogenesis. *Science*, **348**, 812-817.
- Monteys, A. M., Spengler, R. M., Wan, J., Tecedor, L., Lennox, K. A., Xing, Y., *et al.* (2010) Structure and activity of putative intronic miRNA promoters. *RNA*, **16**, 495-505.
- Moore, P. B. & Steitz, T. A. (2011) The roles of RNA in the synthesis of protein. *Cold Spring Harb Perspect Biol*, **3**, a003780.
- Muro, E. M. & Andrade-Navarro, M. A. (2010) Pseudogenes as an alternative source of natural antisense transcripts. *BMC Evol Biol*, **10**, 338.
- Nagano, T., Mitchell, J. A., Sanz, L. A., Pauler, F. M., Ferguson-Smith, A. C., Feil, R., *et al.* (2008) The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science*, **322**, 1717-1720.
- Nakagawa, T., Endo, H., Yokoyama, M., Abe, J., Tamai, K., Tanaka, N., *et al.* (2013) Large noncoding RNA HOTAIR enhances aggressive biological behavior and is associated with short disease-free survival in human non-small cell lung cancer. *Biochem Biophys Res Commun*, **436**, 319-324.
- Nakaya, H. I., Amaral, P. P., Louro, R., Lopes, A., Fachel, A. A., Moreira, Y. B., *et al.* (2007) Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol*, **8**, R43.
- Nan, X., Campoy, F. J. & Bird, A. (1997) MeCP2 is a transcriptional repressor with abundant binding sites in genomic chromatin. *Cell*, **88**, 471-481.
- Napoli, C., Lemieux, C. & Jorgensen, R. (1990) Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell*, **2**, 279-289.
- Neil, H., Malabat, C., d'Aubenton-Carafa, Y., Xu, Z., Steinmetz, L. M. & Jacquier, A. (2009) Widespread bidirectional promoters are the major source of cryptic transcripts in yeast. *Nature*, **457**, 1038-1042.
- Nejepinska, J., Malik, R., Filkowski, J., Flemr, M., Filipowicz, W. & Svoboda, P. (2012) dsRNA expression in the mouse elicits RNAi in oocytes and low adenosine deamination in somatic cells. *Nucleic Acids Res*, **40**, 399-413.
- Nejepinska, J., Malik, R., Moravec, M. & Svoboda, P. (2012) Deep sequencing reveals complex spurious transcription from transiently transfected plasmids. *PLoS One*, **7**, e43283.





Nejepinska, J., Malik, R., Wagner, S. & Svoboda, P. (2014) Reporters transiently transfected into mammalian cells are highly sensitive to translational repression induced by dsRNA expression. *PLoS One*, **9**, e87517.

Nickerson, J. A., Krochmalnic, G., Wan, K. M. & Penman, S. (1989) Chromatin architecture and nuclear RNA. *Proc Natl Acad Sci U S A*, **86**, 177-181.

Nishida, K. M., Okada, T. N., Kawamura, T., Mituyama, T., Kawamura, Y., Inagaki, S., *et al.* (2009) Functional involvement of Tudor and dPRMT5 in the piRNA processing pathway in *Drosophila* germlines. *EMBO J*, **28**, 3820-3831.

Nishikura K, Yoo C, Kim U, Murray JM, Estes PA, Cash FE, *et al.* (1991) Substrate specificity of the dsRNA unwinding/modifying activity. *EMBO J*, **10**, 3523-32.

Nishikura, Kazuko (2010). "Functions and Regulation of RNA Editing by ADAR Deaminases". *Annual Review of Biochemistry*, **79**, 321-349

Nishimasu, H., Ishizu, H., Saito, K., Fukuhara, S., Kamatani, M. K., Bonnefond, L., *et al.* (2012) Structure and function of Zucchini endoribonuclease in piRNA biogenesis. *Nature*, **491**, 284-287.

Noland, C. L. & Doudna, J. A. (2013) Multiple sensors ensure guide strand selection in human RNAi pathways. *RNA*, **19**, 639-648.

Noordermeer, D., Leleu, M., Schorderet, P., Joye, E., Chabaud, F. & Duboule, D. (2014) Temporal dynamics and developmental memory of 3D chromatin architecture at Hox gene loci. *Elife*, **3**, e02557.

O'Connell, MA. Keller, W. (1994) Purification and properties of double-stranded RNA-specific adenosine deaminase from calf thymus. *Proc Natl Acad Sci USA*, **91**, 10596-10600

O'Connell, MA. Gerver, A. Keller, W. (1997) Purification of human double-stranded RNA-specific editase 1 (Hred1) involved in editing of brain glutamate receptor B pre-mRNA. *J Biol Chem*, **272**, 473-478.

Okada, C., Yamashita, E., Lee, S. J., Shibata, S., Katahira, J., Nakagawa, A., *et al.* (2009) A high-resolution structure of the pre-microRNA nuclear export machinery. *Science*, **326**, 1275-1279.

Olivieri, D., Sykora, M. M., Sachidanandam, R., Mechtler, K. & Brennecke, J. (2010) An in vivo RNAi assay identifies major genetic and cellular requirements for primary piRNA biogenesis in *Drosophila*. *EMBO J*, **29**, 3301-3317.

Olsen, P. H. & Ambros, V. (1999) The lin-4 regulatory RNA controls developmental timing in *Caenorhabditis elegans* by blocking LIN-14 protein synthesis after the initiation of translation. *Dev Biol*, **216**, 671-680.

Orom, U. A., Nielsen, F. C. & Lund, A. H. (2008) MicroRNA-10a binds the 5'UTR of ribosomal protein mRNAs and enhances their translation. *Mol Cell*, **30**, 460-471.

Osborne, C. S., Chakalova, L., Brown, K. E., Carter, D., Horton, A., Debrand, E., *et al.* (2004) Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*, **36**, 1065-1071.



Osenberg, S. Paz Yaacov, N. Safran, M. Moshkovitz, S. Shtrichman, R. *et al.* (2010) Alu sequences in undifferentiated human embryonic stem cells display high levels of A-to-I RNA editing. *PLoS One*, **5**, e11173.

Ozsolak, F., Kapranov, P., Foissac, S., Kim, S. W., Fishilevich, E., Monaghan, A. P., *et al.* (2010) Comprehensive polyadenylation site maps in yeast and human reveal pervasive alternative polyadenylation. *Cell*, **143**, 1018-1029.

Pan, Q., Shai, O., Lee, L. J., Frey, B. J. & Blencowe, B. J. (2008) Deep surveying of alternative splicing complexity in the human transcriptome by high-throughput sequencing. *Nat Genet*, **40**, 1413-1415.

Parker, J. S. & Barford, D. (2006) Argonaute: A scaffold for the function of short regulatory RNAs. *Trends Biochem Sci*, **31**, 622-630.

Pasquinelli, A. E., Reinhart, B. J., Slack, F., Martindale, M. Q., Kuroda, M. I., Maller, B., *et al.* (2000) Conservation of the sequence and temporal expression of let-7 heterochronic regulatory RNA. *Nature*, **408**, 86-89.

Patterson, JB. Samuel, CE. (1995) Expression and regulation by interferon of a double-stranded-RNA-specific adenosine deaminase from human cells: evidence for two forms of the deaminase, *Mol. Cell. Biol*, **15**, 5376-5388.

Peisley, A., Lin, C., Wu, B., Orme-Johnson, M., Liu, M., Walz, T., *et al.* (2011) Cooperative assembly and dynamic disassembly of MDA5 filaments for viral dsRNA recognition. *Proc Natl Acad Sci U S A*, **108**, 21010-21015.

Pichlmair, A., Schulz, O., Tan, C. P., Naslund, T. I., Liljestrom, P., Weber, F., *et al.* (2006) RIG-I-mediated antiviral responses to single-stranded RNA bearing 5'-phosphates. *Science*, **314**, 997-1001.

Plath, K., Fang, J., Mlynarczyk-Evans, S. K., Cao, R., Worringer, K. A., Wang, H., *et al.* (2003) Role of histone H3 lysine 27 methylation in X inactivation. *Science*, **300**, 131-135.

Poliseno, L., Salmena, L., Zhang, J., Carver, B., Haveman, W. J. & Pandolfi, P. P. (2010) A coding-independent function of gene and pseudogene mRNAs regulates tumour biology. *Nature*, **465**, 1033-1038.

Ponting, C. P., Oliver, P. L. & Reik, W. (2009) Evolution and functions of long noncoding RNAs. *Cell*, **136**, 629-641.

Portal, M. M., Pavet, V., Erb, C. & Gronemeyer, H. (2015) Human cells contain natural double-stranded RNAs with potential regulatory functions. *Nat Struct Mol Biol*, **22**, 89-97.

Raitskin, O. Cho, D. S, Sperling, J. Nishikura, K and Sperling, R. (2001). RNA editing activity is associated with splicing factors in InRNP particles: The nuclear pre-mRNA processing machinery. *Proc. Natl. Acad. Sci. USA*, **98**, 6571-6576.

Rajasethupathy, P., Antonov, I., Sheridan, R., Frey, S., Sander, C., Tuschl, T., *et al.* (2012) A role for neuronal piRNAs in the epigenetic control of memory-related synaptic plasticity. *Cell*, **149**, 693-707.



- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., *et al.* (2000) The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, **403**, 901-906.
- Rinn, J. L., Kertesz, M., Wang, J. K., Squazzo, S. L., Xu, X., Brugmann, S. A., *et al.* (2007) Functional demarcation of active and silent chromatin domains in human HOX loci by noncoding RNAs. *Cell*, **129**, 1311-1323.
- Rouget, C., Papin, C., Boureux, A., Meunier, A. C., Franco, B., Robine, N., *et al.* (2010) Maternal mRNA deadenylation and decay by the piRNA pathway in the early *Drosophila* embryo. *Nature*, **467**, 1128-1132.
- Saini, H. K., Griffiths-Jones, S. & Enright, A. J. (2007) Genomic analysis of human microRNA transcripts. *Proc Natl Acad Sci U S A*, **104**, 17719-17724.
- Saito, K., Ishizu, H., Komai, M., Kotani, H., Kawamura, Y., Nishida, K. M., *et al.* (2010) Roles for the Yb body components Armitage and Yb in primary piRNA biogenesis in *Drosophila*. *Genes Dev*, **24**, 2493-2498.
- Saito, K., Nishida, K. M., Mori, T., Kawamura, Y., Miyoshi, K., Nagami, T., *et al.* (2006) Specific association of Piwi with rasiRNAs derived from retrotransposon and heterochromatic regions in the *Drosophila* genome. *Genes Dev*, **20**, 2214-2222.
- Saito, K. & Siomi, M. C. (2010) Small RNA-mediated quiescence of transposable elements in animals. *Dev Cell*, **19**, 687-697.
- Sansam, CL. Wells, KS. Emeson, RB. (2003) Modulation of RNA editing by functional nucleolar sequestration of ADAR2. *Proc Natl Acad Sci US*, **100**, 14018-14023.
- Schmitz, K. M., Mayer, C., Postepska, A. & Grummt, I. (2010) Interaction of noncoding RNA with the rDNA promoter mediates recruitment of DNMT3b and silencing of rRNA genes. *Genes Dev*, **24**, 2264-2269.
- Schneider, C., King, R. M. & Philipson, L. (1988) Genes specifically expressed at growth arrest of mammalian cells. *Cell*, **54**, 787-793.
- Seeburg, PH. Hartner, J. (2003) Regulation of ion channel/neurotransmitter receptor function by RNA editing. *Curr. Opin. Neurobiol*, **13**, 279-283.
- Seidl, C. I., Stricker, S. H. & Barlow, D. P. (2006) The imprinted Air ncRNA is an atypical RNAPII transcript that evades splicing and escapes nuclear export. *EMBO J*, **25**, 3565-3575.
- Sen, G. L. & Blau, H. M. (2005) Argonaute 2/RISC resides in sites of mammalian mRNA decay known as cytoplasmic bodies. *Nat Cell Biol*, **7**, 633-636.
- Sen, G. L., Wehrman, T. S. & Blau, H. M. (2005) mRNA translation is not a prerequisite for small interfering RNA-mediated mRNA cleavage. *Differentiation*, **73**, 287-293.
- Serman, A., Le Roy, F., Aigueperse, C., Kress, M., Dautry, F. & Weil, D. (2007) GW body disassembly triggered by siRNAs independently of their silencing activity. *Nucleic Acids Res*, **35**, 4715-4727.
- Sheth, U. & Parker, R. (2003) Decapping and decay of messenger RNA occur in cytoplasmic processing bodies. *Science*, **300**, 805-808.





- Shi, Y. & Manley, J. L. (2007) A complex signaling pathway regulates SRp38 phosphorylation and pre-mRNA splicing in response to heat shock. *Mol Cell*, **28**, 79-90.
- Shinagawa, T. & Ishii, S. (2003) Generation of Ski-knockdown mice by expressing a long double-strand RNA from an RNA polymerase II promoter. *Genes Dev*, **17**, 1340-1345.
- Siomi, M. C., Mannen, T. & Siomi, H. (2010) How does the royal family of Tudor rule the PIWI-interacting RNA pathway? *Genes Dev*, **24**, 636-646.
- Sleutels, F., Zwart, R. & Barlow, D. P. (2002) The non-coding Air RNA is required for silencing autosomal imprinted genes. *Nature*, **415**, 810-813.
- Smith, C. M. & Steitz, J. A. (1998) Classification of gas5 as a multi-small-nucleolar-RNA (snoRNA) host gene and a member of the 5'-terminal oligopyrimidine gene family reveals common features of snoRNA host genes. *Mol Cell Biol*, **18**, 6897-6909.
- Sone, M., Hayashi, T., Tarui, H., Agata, K., Takeichi, M. & Nakagawa, S. (2007) The mRNA-like noncoding RNA Gomafu constitutes a novel nuclear domain in a subset of neurons. *J Cell Sci*, **120**, 2498-2506.
- Song, J. J., Smith, S. K., Hannon, G. J. & Joshua-Tor, L. (2004) Crystal structure of Argonaute and its implications for RISC slicer activity. *Science*, **305**, 1434-1437.
- Soruco, M. M., Chery, J., Bishop, E. P., Siggers, T., Tolstorukov, M. Y., Leydon, A. R., *et al.* (2013) The CLAMP protein links the MSL complex to the X chromosome during Drosophila dosage compensation. *Genes Dev*, **27**, 1551-1556.
- St Laurent, G., Savva, Y. A. & Kapranov, P. (2012) Dark matter RNA: an intelligent scaffold for the dynamic regulation of the nuclear information landscape. *Front Genet*, **3**, 57.
- Struhl, K. (2007) Transcriptional noise and the fidelity of initiation by RNA polymerase II. *Nat Struct Mol Biol*, **14**, 103-105.
- Svoboda, P., Stein, P., Hayashi, H. & Schultz, R. M. (2000) Selective reduction of dormant maternal mRNAs in mouse oocytes by RNA interference. *Development*, **127**, 4147-4156.
- Tam, O. H., Aravin, A. A., Stein, P., Girard, A., Murchison, E. P., Cheloufi, S., *et al.* (2008) Pseudogene-derived small interfering RNAs regulate gene expression in mouse oocytes. *Nature*, **453**, 534-538.
- Tay, Y., Kats, L., Salmena, L., Weiss, D., Tan, S. M., Ala, U., *et al.* (2011) Coding-independent regulation of the tumor suppressor PTEN by competing endogenous mRNAs. *Cell*, **147**, 344-357.
- Terenzi, F., deVeer, M. J., Ying, H., Restifo, N. P., Williams, B. R. & Silverman, R. H. (1999) The antiviral enzymes PKR and RNase L suppress gene expression from viral and non-viral based vectors. *Nucleic Acids Res*, **27**, 4369-4375.
- Till, S., Lejeune, E., Thermann, R., Bortfeld, M., Hothorn, M., Enderle, D., *et al.* (2007) A conserved motif in Argonaute-interacting proteins mediates functional interactions through the Argonaute PIWI domain. *Nat Struct Mol Biol*, **14**, 897-903.
- Tripathi, V., Ellis, J. D., Shen, Z., Song, D. Y., Pan, Q., Watt, A. T., *et al.* (2010) The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*, **39**, 925-938.



- Tsai, M. C., Manor, O., Wan, Y., Mosammaparast, N., Wang, J. K., Lan, F., *et al.* (2010) Long noncoding RNA as modular scaffold of histone modification complexes. *Science*, **329**, 689-693.
- Tsuiji, H., Yoshimoto, R., Hasegawa, Y., Furuno, M., Yoshida, M. & Nakagawa, S. (2011) Competition between a noncoding exon and introns: Gomafu contains tandem UACUAAC repeats and associates with splicing factor-1. *Genes Cells*, **16**, 479-490.
- Tufarelli, C., Stanley, J. A., Garrick, D., Sharpe, J. A., Ayyub, H., Wood, W. G., *et al.* (2003) Transcription of antisense RNA leading to gene silencing and methylation as a novel cause of human genetic disease. *Nat Genet*, **34**, 157-165.
- Vagin, V. V., Sigova, A., Li, C., Seitz, H., Gvozdev, V. & Zamore, P. D. (2006) A distinct small RNA pathway silences selfish genetic elements in the germline. *Science*, **313**, 320-324.
- Vanhee-Brossollet, C. & Vaquero, C. (1998) Do natural antisense transcripts make sense in eukaryotes? *Gene*, **211**, 1-9.
- Vasileva, A., Tiedau, D., Firooznia, A., Muller-Reichert, T. & Jessberger, R. (2009) Tdrd6 is required for spermiogenesis, chromatoid body architecture, and regulation of miRNA expression. *Curr Biol*, **19**, 630-639.
- Vassar, R., Bennett, B. D., Babu-Khan, S., Kahn, S., Mendiaz, E. A., Denis, P., *et al.* (1999) Beta-secretase cleavage of Alzheimer's amyloid precursor protein by the transmembrane aspartic protease BACE. *Science*, **286**, 735-741.
- Vaucheret, H., Beclin, C. & Fagard, M. (2001) Post-transcriptional gene silencing in plants. *J Cell Sci*, **114**, 3083-3091.
- Wang, Z. & Burge, C. B. (2008) Splicing regulation: from a parts list of regulatory elements to an integrated splicing code. *RNA*, **14**, 802-813.
- Watanabe, T., Takeda, A., Tsukiyama, T., Mise, K., Okuno, T., Sasaki, H., *et al.* (2006) Identification and characterization of two novel classes of small RNAs in the mouse germline: retrotransposon-derived siRNAs in oocytes and germline small RNAs in testes. *Genes Dev*, **20**, 1732-1743.
- Watanabe, T., Tomizawa, S., Mitsuya, K., Totoki, Y., Yamamoto, Y., Kuramochi-Miyagawa, S., *et al.* (2011) Role for piRNAs and noncoding RNA in de novo DNA methylation of the imprinted mouse Rasgrf1 locus. *Science*, **332**, 848-852.
- Watanabe, T., Totoki, Y., Toyoda, A., Kaneda, M., Kuramochi-Miyagawa, S., Obata, Y., *et al.* (2008) Endogenous siRNAs from naturally formed dsRNAs regulate transcripts in mouse oocytes. *Nature*, **453**, 539-543.
- Webster, A., Li, S., Hur, J. K., Wachsmuth, M., Bois, J. S., Perkins, E. M., *et al.* (2015) Aub and Ago3 Are Recruited to Nuage through Two Mechanisms to Form a Ping-Pong Complex Assembled by Krimper. *Mol Cell*, **59**, 564-575.
- Werner, A., Cockell, S., Falconer, J., Carlile, M., Alnumeir, S. & Robinson, J. (2014) Contribution of natural antisense transcription to an endogenous siRNA signature in human cells. *BMC Genomics*, **15**, 19.



- Wickham, L., Duchaine, T., Luo, M., Nabi, I. R. & DesGroseillers, L. (1999) Mammalian staufen is a double-stranded-RNA- and tubulin-binding protein which localizes to the rough endoplasmic reticulum. *Mol Cell Biol*, **19**, 2220-2230.
- Wightman, B., Ha, I. & Ruvkun, G. (1993) Posttranscriptional regulation of the heterochronic gene *lin-14* by *lin-4* mediates temporal pattern formation in *C. elegans*. *Cell*, **75**, 855-862.
- Woo, C. J. & Kingston, R. E. (2007) HOTAIR lifts noncoding RNAs to new levels. *Cell*, **129**, 1257-1259.
- Wu, L., Fan, J. & Belasco, J. G. (2006) MicroRNAs direct rapid deadenylation of mRNA. *Proc Natl Acad Sci U S A*, **103**, 4034-4039.
- Wu, L., Murat, P., Matak-Vinkovic, D., Murrell, A. & Balasubramanian, S. (2013) Binding interactions between long noncoding RNA HOTAIR and PRC2 proteins. *Biochemistry*, **52**, 9519-9527.
- Wutz, A. & Jaenisch, R. (2000) A shift from reversible to irreversible X inactivation is triggered during ES cell differentiation. *Mol Cell*, **5**, 695-705.
- Xu, H., He, X., Zheng, H., Huang, L. J., Hou, F., Yu, Z., *et al.* (2014) Structural basis for the prion-like MAVS filaments in antiviral innate immunity. *Elife*, **3**, e01489.
- Yap, K. L., Li, S., Munoz-Cabello, A. M., Raguz, S., Zeng, L., Mujtaba, S., *et al.* (2010) Molecular interplay of the noncoding RNA ANRIL and methylated histone H3 lysine 27 by polycomb CBX7 in transcriptional silencing of INK4a. *Mol Cell*, **38**, 662-674.
- Yekta, S., Shih, I. H. & Bartel, D. P. (2004) MicroRNA-directed cleavage of HOXB8 mRNA. *Science*, **304**, 594-596.
- Yi, R., Qin, Y., Macara, I. G. & Cullen, B. R. (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev*, **17**, 3011-3016.
- Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., *et al.* (2004) The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. *Nat Immunol*, **5**, 730-737.
- Yoneyama, M., Onomoto, K., Jogi, M., Akaboshi, T. & Fujita, T. (2015) Viral RNA detection by RIG-I-like receptors. *Curr Opin Immunol*, **32**, 48-53.
- Yoon, J. H., Abdelmohsen, K., Srikantan, S., Yang, X., Martindale, J. L., De, S., *et al.* (2012) LincRNA-p21 suppresses target mRNA translation. *Mol Cell*, **47**, 648-655.
- Young, L. E., Sanduja, S., Bemis-Standoli, K., Pena, E. A., Price, R. L. & Dixon, D. A. (2009) The mRNA binding proteins HuR and tristetraprolin regulate cyclooxygenase 2 expression during colon carcinogenesis. *Gastroenterology*, **136**, 1669-1679.
- Yu, W., Gius, D., Onyango, P., Muldoon-Jacobs, K., Karp, J., Feinberg, A. P., *et al.* (2008) Epigenetic silencing of tumour suppressor gene p15 by its antisense RNA. *Nature*, **451**, 202-206.
- Zhao, J., Sun, B. K., Erwin, J. A., Song, J. J. & Lee, J. T. (2008) Polycomb proteins targeted by a short repeat RNA to the mouse X chromosome. *Science*, **322**, 750-756.



Zhong, X. Y., Ding, J. H., Adams, J. A., Ghosh, G. & Fu, X. D. (2009) Regulation of SR protein phosphorylation and alternative splicing by modulating kinetic interactions of SRPK1 with molecular chaperones. *Genes Dev*, **23**, 482-495.

